The Dissertation Committee for Gregory Belknap Cumpton Certifies that this is the approved version of the following dissertation:

The Effect of the Policy of Reconstitution on Student Achievement in Texas

Committee:

Jane Lincove, Supervisor

Chris King

Pedro Reyes

Chandler Stolp

Uri Triesman

The Effect of the Policy of Reconstitution on Student Achievement in Texas

by

Gregory Belknap Cumpton, B.S., M.P.Aff.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2015

Dedication

No one possessed more constant and openly expressed faith in me than my beloved wife, Cheryl Ann Goveia. Had I only relied on my own minimal faith in my capacity and effort to complete this work, I would never have finished. Cheryl listened to all my many doubts but never wavered in her unending confidence that I would complete this work. More than providing emotional and intellectual support to this work, Cheryl's experiences teaching art at Pearce Middle School in Austin, TX during one of the years the school reconstituted led, through many fits and starts, to the topic of this dissertation. Her insight into this reform greatly enhanced this work.

The Effect of the Policy of Reconstitution on Student Achievement in Texas

Gregory Belknap Cumpton, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Jane Lincove

The failure of schools across the country to ensure students meet federal, state, and community standards continuously plagues the education system. More than a quarter of all schools in the nation failed to meet federal requirements in 2007, with 38% failing to do so in 2010. By 2011 that figure rose to nearly 50%. Failing schools ostensibly produce failing students who experience poorer outcomes than their peers including reduced earnings over their lifetime. A potential solution to failing schools is to reconstitute them. School reconstitution requires all staff at a failing school to reapply for their positions with the stated aim of improving student achievement. Started as a court-mandated desegregation action in San Francisco in 1983, school reconstitution quickly spread across the country in the 1990s. Incorporated into local and state accountability systems, scholars estimate thousands of schools reconstituted between 1983 and 2011. Despite its prevalence, information regarding how reconstitution began, spread, and made its way into Texas statute is scarce and theories related to why reconstitution should improve student performance lack cohesion. Even worse, little to no quantitative evidence demonstrates whether reconstitution improves student achievement. This dissertation takes advantage of a Texas law passed in 2003 mandating that schools failing to meet state standards for two years in a row must reconstitute. Estimated effects of reconstitution on student achievement apply state-wide student and school data between 2003 and

2011. Several methods, including regression discontinuity and student-level fixed-effects determine whether reconstitution improves student achievement and if developed theories explain this improvement. Discussion includes national, state, and local policy recommendations.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The failure of schools across the country to ensure students meet federal, state, and community standards continuously plagues the education system. Passage of No Child Left Behind (NCLB) in 2001, which mandates that schools must make adequate yearly progress (AYP) toward state-defined student academic proficiency goals, hints at the scale and longevity of this issue. More than a quarter of all schools in the nation failed to make AYP in 2007, with 38% failing to do so in 2010 (Usher 2011). By 2011 that figure rose to nearly 50% (Usher 2011). In 12 states, half or more of schools failed to meet AYP in 2010 (Usher 2011). Naturally, considerable public discourse exists around the need to improve failing schools from the late 1990s to the present.

"We cannot and must not tolerate failing schools. We need to stop making excuses and get on with the business of fixing our schools," Secretary of Education, Richard Riley, Feb 18th, 1997 through (Doherty and Abernathy 1998).

"We should get out of the business of labeling schools as failures and create a new law that is…focused on the schools and students most at risk," Arne Duncan (DoE 2011).

Why should failing schools concern us? Besides the moral imperative of providing an education to its population, concerns about failing schools center on their pre-condition: failing students. Failing schools include measurable sizes of students demonstrating poor academic performance, a prerequisite for their classification as failing. Research over the last forty years shows that poor academic performance is associated with a wide range of detrimental outcomes. If failing schools lead to poor academic preparation, the aggregate effects of individual poor performance likely contribute to larger societal ills.

Poor academic performance in public school increases the probability of dropping out of high school, which is itself associated with longer-term detrimental outcomes

throughout an individual's life (Hoffmann, Erickson et al. 2013). For example, an individual's highest level of education broadly determines their long-term earnings potential; those with at least a college degree earn much more over their lifetime than those with less education. In 2014, working high school dropouts earned almost half as much each week as the national average (BLS 2015). Other types of work outcomes, including unemployment and long-term income stability, favor the educated. In 2014, the unemployment rate of individuals without a high school diploma was 9% compared to the national average of 5%, a consistent trend for decades (BLS 2015). Workers who did not graduate from high school were dramatically impacted by the Great Recession, with rates of unemployment rising to nearly 16% for individuals in this group (BLS 2015).

Poor individual academic performance in elementary, middle, and high school also reduces the probability of enrolling in college, even if students do graduate from high school, which may limit future job choices and also earnings. Additionally, students who do not complete college engage in unhealthy activities at a higher rate. Nationally, 22% of adults whose highest education was a high school degree are smokers, compared to roughly 9% of adults with a college degree (CDC 2015). More than a third of individuals between the ages of 25 and 34 whose highest education level was a high school diploma or less were classified as obese, while this is true for just 1 in 5 individuals from the same age group who obtained a bachelor's degree or higher (Baum, Ma et al. 2013).

Lower levels of the highest education earned are associated with an individual's positive involvement in their community. A slim majority of individuals between the ages of 25 and 44 with some college voted in the 2004 presidential election, while 76% of college graduates did so. Among individuals without a high school diploma, the voting rate was just 27% (Baum, Ma et al. 2013). Individuals with a high school diploma or less are notably less likely to volunteer their time in their community, with fewer than 2 in 10

doing so; those with a college degree more than double that share of volunteering in their communities (Baum, Ma et al. 2013). Poor academic achievement, regardless of high school graduation status, is associated with future delinquent behavior (Hoffmann, Erickson et al. 2013). Students who do not finish high school are more likely to experience incarceration or juvenile detention (Dillon 2009).

Thus, students who are not academically prepared through the public education system likely cost the US economy billions of dollars of consumer spending, lost tax revenue, and increased needs for state support and resources over their lifetimes.

While placing the blame for all of these outcomes on failing public schools is surely a stretch, they certainly do not help. If formerly failing schools provide supports and learning to students such that they leave each grade knowledge-ready for the next, students might be less likely to drop out of high school and thus avoid the negative consequences of the long-term reduction in their human capital potential. Thus, states have an inherent interest in ensuring that schools are successful in their efforts at teaching students. Schools readily identifiable as failing state standards potentially place long-term burdens on students, and, in turn, their families, their cities, counties, and the state.

There is continuing debate about what to do with failing schools: should we replace failing public schools with charters, or have the state take over failing schools, or provide vouchers for private schools? One potential solution to the problem of failing schools is reconstitution, which requires removal of all or most of a school's staff to improve student academic performance (Malen, Croninger et al. 2002). Reconstitution has been applied to schools across the country since 1983 (Ruiz-de-Velasco, 1999). Beginning in 2003 in Texas, schools that failed to meet state standards for two years in a row are required to reconstitute (SB 618 2003). Over 135 schools with over 58,000 students have experienced reconstitution in Texas alone.

Despite its widespread application, little to no work traces how reconstitution came to be included in Texas statute, whether and to what extend it has been applied, and whether reconstitution truly produces the intended effect on student achievement. This dissertation answers the following research questions:

1. Where did reconstitution come from and how did it become a feature of state-wide education policy in Texas?
2. How is reconstitution applied in Texas?
3. Does reconstitution help failing schools improve their minimum passing score?
4. What is the effect of reconstitution on student performance on mathematics and reading tests?

The second chapter provides an overview of how school reconstitution came to be included in Texas' education accountability system using the multiple streams framework of Kingdon (Kingdon 1984). Reconstitution benefited from three distinct policy windows where the streams of politics and problems became linked to reconstitution, leading to its initial adoption in San Francisco, its diffusion beyond the borders of California, to its inclusion in law in Texas.

The third chapter examines and synthesizes the theories associated with reconstitution. Among the reform ideas available for failing schools, reconstitution may appear counterintuitive. If voluntary teacher turnover inhibits student achievement (Ronfeldt, Loeb et al. 2013), how can involuntary teacher turnover improve student achievement? The second chapter reviews prior literature on school reconstitution, much of which uses qualitative techniques to describe the potentially disruptive nature of reconstitution on teachers and, occasionally, on students.

The fourth chapter discusses the data available to determine whether school reconstitution actually improves student outcomes and provides a description of the methods applied to investigate the effect of this policy on student achievement. A number of methods are applied, including a school-level regression discontinuity design and student-level fixed effects models. This section also discusses the state accountability system in some detail, as this is the means by which schools are identified for reconstitution.

The fifth chapter examines the results of the analyses described in the third chapter. This chapter looks first for evidence of staff turnover at reconstituted schools. This chapter then examines the effect of reconstitution on student achievement in Texas, assessing whether the policy improved school-level performance and whether students who experienced reconstitution improved their performance on state-mandated tests.

The sixth chapter discusses the results of this study and makes national, state, district, and school policy recommendations related to reconstitution specifically as well as discusses attempts to study and improve school-level performance more broadly.

Chapter 2: How School Reconstitution Came to Texas

Reconstitution follows a strange path from a court consent decree in a liberal city in a liberal state on the west coast to reside in codified law in politically conservative Texas. This chapter follows the evolution and diffusion of reconstitution from its origin to its adoption as a mandatory school sanction policy in Texas. The complexity inherent in this path requires applying a model of how the policy process works. John W. Kingdon indicated that "[t]he separate streams of problems, policies, and politics come together at certain critical times. Solutions become tied to problems, and both of them are joined to favorable political forces. This coupling is most likely when policy windows--opportunities for pushing pet proposals or conceptions of problems--are open." (Kingdon 1984; pg. 20). Reconstitution successfully traverses through three such policy windows. The first occurs in 1983 in San Francisco, when reconstitution addresses the problem of school segregation. During this period, courts monitor the creation and implementation of reconstitution. The second policy window occurs in 1992 in San Francisco when a modified form of reconstitution, now tied to school level academic standards, becomes district policy due in part to the advocacy of the districts' superintendent. The third window occurs in 2003 when Texas adopts reconstitution as a solution for failing schools.

THE FIRST POLICY WINDOW

In his book, "Agendas, Alternatives, and Public Policies," Kingdon proposes that the policy process focuses on "problems, policies, and politics" (Kingdon, 1984; pg. 87). Kingdon's framework elucidates how agencies adopt certain policies, positing that when these three streams meet they create a window of opportunity for policy adoption.[1]

[1] There is some controversy in the wider literature at applying Kingdon's streams (which were developed for the purpose of analyzing federal policies) to state policies. This particular analysis stands clearly on the side that Kingdon's streams framework remains relevant in many contexts, not merely state policies but also court and school district policies.

Kingdon's streams framework quite adequately describes the history of the adoption of reconstitution policies from their creation in 1983 to codification as a mandatory sanction in Texas twenty years later. Part of the reason for the popularity of employing Kingdon's multiple streams framework in education policy stems from the chaotic and distributive nature of education policy in general, which other policy frameworks cannot adequately describe (Lewis 2013).

Identifying the Problem: School Segregation and Integration

School reconstitution begins as a desegregation court case, but due to constraints on federal court actions and shifts in national perspectives on education, district officials and the courts develop new tools to ensure integration of students and teachers. Reconstitution, with its emphasis on replacing existing staff with highly qualified teachers, serves as a compromise to meet those needs.

Desegregation

Thus, the deep origins of reconstitution lie with Brown v. Board of Education, which pronounces the fundamental principle that "racial discrimination in public education is unconstitutional [and that] all provisions of federal, state, or local law requiring or permitting such discrimination must yield to this principle," (1954, pg. 298). Though the Supreme Court clearly overturns previous legal cases establishing and allowing the maintenance of lawful discrimination, the decision does not outline a clear path to remedy existent discrimination (Daugherity and Bolton 2008). The last line in the first paragraph of Chief Justice Warren's opinion indicates that, "there remains for consideration the manner in which relief is to be accorded," (1954). The long and arguably arduous process of integration across the country lasts for decades, with implementation of integration often falling on federal courts (Kluger 2004). Exactly when a school or school district integrates

varies based on a number of factors, including the clarity of the discrimination, the local impetus to eliminate segregation, and whether discriminatory practices are brought to the federal Justice Department or lawsuits brought against the school district or state (Kluger 2004).

In San Francisco Unified School District (SFUSD), as in other school districts, multiple court cases over many years alleging discrimination eventually led to several court-mandated integration efforts (Ruiz-de-Velasco 1999). As early as 1971, Johnson v. SFUSD found *de jure* discrimination, indicating that the school district maintained the segregation of its schools through established written policies (Consent Decree 1983, fn. #4). SFUSD appealed this decision, and a court case in 1974 found in favor of the district. That court ruling, however, still "directed [the] District to continue to comply with the desegregation plan ordered by the [original] trial court," (Consent Decree 1983, fn. #4). The district responded to these lawsuits in the early 1970s by introducing bussing within the district in an effort to meet court-mandated requirements, though some parents moved to the suburbs or transferred their children to suburban districts to avoid integration, engaging in a practice referred to as 'white flight'.[2] Despite the court directive and efforts by the district, policies and practices resulted in the continued segregation of schools (Rojas 1996).

A 1978 lawsuit by the National Association for the Advancement of Colored People (NAACP) against SFUSD, which would come to spawn school reconstitution, addressed the long existent segregation within the school district (Consent Decree 1983). The NAACP contended that the school district ensured students attended segregated schools by following specific policies which constructed a dual school system: one for minority

[2] According to the Merriam Webster dictionary the first use of this term occurred in 1967.

students and another for non-minority students (Consent Decree 1983, fn. #4). In a memorandum to the court in 1980, the NAACP indicated that the long term patterns of segregation in San Francisco originated as far back as the 1850s. Notable circumstantial evidence of recent district collusion in this segregation included the following:

> Constructing new schools and annexes, leasing private property for school use, and utilizing portable classrooms in order to incorporate extant residential segregation into the District; establishing feeder patterns, transfer and reassignment policies, optional and mandatory attendance zones to situate children in racially isolated schools; implementing racially discriminatory testing procedures, disciplinary policies, and tracking systems within schools and classrooms; hiring and assigning faculty and administrative personnel, and allocating financial resources in a discriminatory manner. (Consent Decree 1983, §IA).

The court agreed with the NAACP's assessment on the status and cause of segregation within the school district. To preclude court-mandated actions, the NAACP and SFUSD commenced settlement negotiations in 1982 to develop systems and policies to address the district's segregation (Clearinghouse 2015). Eventually these settlement negotiations created the first instances of school reconstitution.

Shift in National Focus

These consent decree negotiations occurred as the national focus on education shifted away from concerns about the inequalities of segregation and toward school quality and school accountability (Orfield 1999). Between 1954 and 1982, federal and Supreme Court cases greatly expanded the authority of courts and the federal government to enforce integration (Kluger 2004). As efforts to integrate spread outside of the south and from students to teachers, reaching into more and more lives, support for the means (e.g., busing) to integrate schools waned (Kluger 2004). Beginning with the Milliken v. Bradley case in 1974, which prevented integration across district boundaries, the expansion of student desegregation efforts across the country slowed (Milliken v. Bradley 1974). Voter

opposition to integration led to California adding Proposition 1 through popular vote to the state constitution in 1979, indicating that state courts could not integrate schools through busing (Crawford v. Los Angeles Board of Education 1982).

Desegregation requirements applied not only to students, but to the entire education system, including teachers (Brown v. Board 1954). Teacher integration proved more complex and more difficult to implement than student integration. In 1966, more than a decade after Brown, Judge John Minor Wisdom noted that "[t]he most difficult problem in the desegregation process is the integration of faculties," (United States v. Jefferson County Board of Education 1966 through Parker 2008, pg 11).[3] Some of the main problems encountered in desegregating teachers stem from the flexibility of teacher labor markets; teachers required to move schools within a school district may choose employment in another nearby district. While little scholarship has focused directly on the issue of desegregating teachers, Parker indicated that "[i]n sum, parents, teachers, and school districts strongly resisted the integration of teachers..." (Parker 2008, pg 14). Despite resistance, as late as 1976, some locales, such as Los Angeles, still bore witness to the expansion of contentious teacher integration efforts (Miller and Montgomery 1976).

Shifts in public opinion about the benefits of desegregation combined with court cases limiting the continued expansion of desegregation beyond district borders would eventually curtail public elementary and secondary desegregation efforts. Scholars assessing the effect of desegregation reached the conclusion that it "caused significant turmoil [and]...resulted in public systems that are even less effective in providing quality education than they once had been," (Biegel 1988). The way forward in education reform, though not yet fully formed, had begun to inexorably move away from using integration as

[3]Judge Wisdom was on the 5th circuit of the court of appeals from 1957-1977.

the means of resolving educational inequalities. By the early 1980s the long great arc of American integration approached the beginning of the end. Though by 1988, "almost 45% of black students in the United States...[would be] attending majority-white schools," this would be the high-water mark of integration in the United States (SPLC 2004).

The same year San Francisco reconstituted schools for the first time (1983), the publication of A Nation at Risk placed national focus on the generally poor state of the education system in America outside of the context of segregation (Gardner 1997). A Nation at Risk's focus on the poor levels of education American students received struck a chord, becoming an intrinsic part of the national consciousness (Zavadsky 2012). The report famously indicated that, "[i]f an unfriendly foreign power had attempted to impose on America the mediocre educational performance that exists today, we might well have viewed it as an act of war," (Gardner 1997, pg. 5). Surely this statement disheartened many of those who had been at the forefront of education reform efforts over the previous thirty years.

A Nation at Risk placed renewed focus on the generally poor state of education and shifted the conversation away from integration as a solution to educational inequalities. The report mostly neglects the recent context race had played in education in America, ignoring how segregated schools were deemed unequal but declaring that

> All, regardless of race or class or economic status, are entitled to a fair chance and to the tools for developing their individual powers of mind and spirit to the utmost. This promise means that all children by virtue of their own efforts, competently guided, can hope to attain the mature and informed judgement (sic) needed to secure gainful employment, and to manage their own lives, thereby serving not only their own interests but also the progress of society itself. (Gardner 1997, pg. 4).

This reference to race without actually referencing race presaged a new, and developing, post-desegregation era of education reform. The fault for terrible schools lay

not with racial isolation, but with poor curriculum, low pay and meager preparation of teachers, a lack of appropriate fiscal supports, and deficiencies of leadership (Gardner 1997). Just one month after the publication of A Nation at Risk, the NAACP and SFUSD agreed to a final consent decree.

Developing and Implementing Consent Decree Policies

The first cases of reconstitution arose out of the court-mandated and monitored negotiations between SFUSD and the NAACP. Kingdon believes it nearly impossible to identify the origin of any policy, noting that anyone can generate and develop policy ideas (Kingdon 1984). He also notes that "there is no logical place to stop" when tracking down the origins of policy ideas (Kingdon 1984, pg. 73). Kingdon describes "tracing origins...to be futile," when it comes to policies (Kingdon 1984, pg. 85). Tracing the origins of policies may yield less relevance for some policy makers and analysts than examining the process of their adoption, but there is still inherent value in tracing these origins. When origins of policies are found, they describe the initial theory and conditions under which the policy was intended to work. While policy ideas may be reworked as they diffuse, the underlying theoretical context should still be present, if hidden, as the policy continues to be implemented in its new form. Too often policy makers and analysts focus on the present application of a policy and neglect this historical context. Many thousands of policies were developed to solve the problems of their times and then were re-worked and manipulated to solve new problems.

The consent decree comprehensively outlined measures necessary for the long-term desegregation of San Francisco Unified School District students by first designating what racial or ethnic groups were recognized by the decree, which demonstrates the incredible diversity of backgrounds within the district: "Spanish-surnamed, Other White, Black,

Chinese, Japanese, Korean, Filipino, American Indian, and other Non-White." (Consent Decree 1983, pg. 37). The decree ordered that, "[r]egular schools are to enroll students from no fewer than four of the nine racial/ethnic groups identified above, and are to ensure that no more than 45 percent of their enrollments are comprised of students from any one race or ethnic group," (Consent Decree 1983, pgs. 41-42). Considering that many of the schools within the district were composed of close to 100% ethnically and racially isolated populations, this remarkable change would take considerable effort to implement.

The court attempted to address some of the underlying causes of this long-term racial isolation, but their current environment, including several sources of segregation beyond the reach of the court, required some inventive solutions. The decree recognized the role of housing policies on school segregation, which produce racially and economically isolated populations. But rather than assume that district measures could fully compensate for this, the consent decree attempted to deal directly with the issue of segregation within public housing. The consent decree tasked an expert in housing policy to analyze housing data and make suggestions aligning housing policies with the wider school desegregation objectives. Recommendations would then be communicated "to the Mayor of San Francisco, the San Francisco Public Housing Authority, the San Francisco Redevelopment Agency, and concerned state and federal agencies," (Consent Decree 1983, pg. 57). The consent decree required that each person and agency respond to these recommendations annually.

Residential segregation patterns extended beyond public housing. As in many school districts across the country, San Francisco experienced the phenomenon of "white flight," where parents either moved to the suburbs or transferred their children to a suburban district. Whereas other courts in other cities in other years extended the reach of bussing efforts into outlying districts, the Milliken v. Bradley case in 1974 specifically

prohibited this action (Milliken v. Bradley 1974). The consent decree, while restricted from preventing families from leaving the district, mandated that "suburban school districts...receive no state aid for students residing in San Francisco whose interdistrict transfers adversely affect San Francisco's school desegregation effort," (Consent Decree 1983, pg. 41). This provision removed any incentive for outlying districts to accept student transfers from SFUSD.

Families of military personnel stationed within San Francisco could enroll their children in private schools rather than nearby local schools. Rather than being isolated cases, so many families enrolled their children in private schools that military bases provided bus transportation between the bases and the private schools. While the court could not mandate public school enrollment, it took issue with the practice of military authorities providing bussing and requested it to stop (Consent Decree 1983, pg. 41).

The decree also recognized the need to integrate school staff, mandating that "the racial/ethnic composition of the full-time faculty, administrators, and other staff of the District reflects that of the District's student enrollment," (Consent Decree 1983, pg. 42). The consent decree took sweeping efforts to tackle multiple sources of segregation in its efforts to fully integrate SFUSD.

The decree also incorporated many other provisions clearly intended to improve academic performance, which was to be used to "evaluate the continued effort to achieve academic excellence throughout the system," (Consent Decree 1983, pg. 42). Student academic achievement included "student test scores and academic results," (Consent Decree 1983, pg. 42). In fact, a required annual report had to "include a section on S.F.U.S.D.'s progress toward the goal of academic excellence," (Consent Decree 1983, pg. 58). There are no provisions outlining what consequences would occur if academic excellence was not reached, nor even a clear definition of academic excellence . This vague

notion of improving student performance found critics in the community. During the period of community comment, some viewed this pursuit of academic excellence as a key, and important, feature of the decree. In response, the court made it very clear that the consent decree was not specifically related to academic performance:

> "This is a school desegregation action…The remedy sought and the remedy proposed is systemwide desegregation. This is not an action to establish an entitlement to a certain standard of academic excellence or to a right to certain programs to meet specific needs," (Consent Decree 1983, pg. 49).

Thus, while academic achievement would be monitored, it was not intended to be the focus of the consent decree and was certainly not established as the underlying purpose of school reconstitution.[4]

The consent decree established an entire framework of reforms to ensure that San Francisco Unified School District would end racial isolation, reaching into the homes of students in public housing, the families of resident students who enrolled their children outside of the district, and the resources of the state of California to ensure the success of this effort. The consent decree mentions reconstitution only three times and ordered reconstitution for only a small number of schools. The first cases of reconstitution on record are intrinsically tied to broader and deeper efforts to end segregation in schools through an environment of court oversight and an expansion of fiscal resources provided by the state. While the primary purpose of the consent decree, desegregating the school district, diminishes greatly over the next thirty years, the singular element of reconstitution lasts much longer.

[4] Rozmus (1998) disputes this assertion, despite the above quoted evidence to the contrary.

School Reconstitution

One region of SFUSD received special attention: the Bayview-Hunters Point area. An area singled out because "[h]istorically, these schools have endured the greatest amount of racial concentration and popular contempt in the District," (Consent Decree 1983, pg. 49). Select schools would be subject to a provision called *reconstitution*: "The SFUSD shall declare all staff and administrative positions in the Bayview-Hunters Point Schools open, and shall reconstitute the staff and administration of those schools *on the basis of a desegregation plan*," (Consent Decree 1983, pg. 55). Thus, reconstituted schools removed old staff and then rehired new staff as part of efforts to desegregate them.

Besides simply stating that schools in the Bayview-Hunter's Point area would need to reconstitute their staff, the decree discussed additional sweeping changes at specific schools and assigned additional fiscal resources to the region including between $1 and $1.5 million in the first year (Consent Decree 1983). Advocates of reconstitution later focused on purported improvements in academic performance at reconstituted schools, but the consent decree mandated specific school-level reforms outside of reconstitution which might also lead to improved academic performance.

These changes included converting some of the schools into magnet schools or schools with specialized curriculum. The Sir Francis Drake School, along with reconstituting, would change "from a basic K-5 elementary school to a computer-assisted instruction and computer science and awareness elementary school," (Consent Decree 1983, pg. 55). The consent decree required that the school worked with a local university to develop the curriculum and that the principal should also teach at that university. To ensure the academic integrity of the school staff at the school would engage in professional development together with college staff.

One school, the Dr. Charles R. Drew School, converted into a middle school. Vertical teams established "between the Drew School and Lowell High School for course development and implementation of courses…would lead to high secondary-level academic performance," (Consent Decree 1983, pg. 55). Furthermore, the catchment area for this school would be citywide and the school would provide transportation to all attendees.

Thus, the application of school reconstitution occurred at some schools whose entire mission changed: teaching different ages of students; increased numbers of counseling staff; pulling students from new and different catchment areas; even defining the purpose of at least one campus to focus on computer science (Consent Decree 1983). Furthermore, reconstituted schools were embedded in a much broader effort to desegregate the entire district including additional resources provided by the state as mandated by the court. For the next nine years, the court intended to monitor the progress of all elements of the consent decree in annual reports (Ruiz-de-Velasco 1999). Recognizing the cost of such measures to the school district, the court ordered that the state would cover all compliance costs (Consent Decree 1983).

But why reconstitute schools? Courts integrated students and teachers in schools for decades without 'reconstituting' them. Perhaps the lack of public enthusiasm surrounding teacher integration led the court to develop alternate means to reach integrated schools. Essentially, reconstitution desegregated both students and teachers and this was done without invoking teacher integration. By having all teachers vacate a campus and require old and new staff re-apply, integration of teachers could occur without mandated district re-assignment of teachers. This avoided political difficulties teacher integration engendered. Thus, the court, school district, and the NAACP developed through the

consent decree a publicly acceptable manner to integrate both SFUSD students and teachers at specific schools.

Thus the federal court overseeing the segregation lawsuit resolution sought to solve the problem of segregated schools but encountered limitations on its authority to demand full and immediate integration for both teachers and students using means such as teacher reassignment and student bussing. Into this politically limiting landscape, consent decree parties developed alternative solutions including reconstitution. This combination of problem, policy solution, and political process opened a policy window for the initial implementation of reconstitution.

THE SECOND POLICY WINDOW

While the origin of the concept of reconstitution began in 1983, a more recognizable version developed in the early 1990s. Kingdon describes the origin of many policies as originating from a "policy primeval soup" (Kingdon 1984, pg. 116). During this time in the soup, "people recombine familiar elements into a new structure or proposal," (Kingdon 1984, pg. 117). Concepts go through a process of changes, combinations, and logical testing. Kingdon argues that "the elements that go into a policy change can be understood better than how they came into contact and were combined with one another," (Kingdon 1984, pg. 139). Unusually, school reconstitution provides a clear example of how these concepts came together as well as who advocated for this policy idea within SFUSD. Reconstitution escaped the small environment of San Francisco and successfully diffused across the country because it met Kingdon's criteria for survival.

Transformation in San Francisco

In the early 1990's the concept of reconstitution, originally tied to integration, combined with early school accountability efforts linked to performance on mandated

academic tests. This bound school reconstitution to the ideas of probation and sanction. Probation, essentially a warning, occurred for schools initially failing set expectations. Sanction, in this case reconstitution, occurred if initially failing schools continued to underperform.

In July of 1992, a committee of court-appointed experts led by Garry Orfield, a professor of education and social policy at Harvard, submitted a court-requested report evaluating the effects of the consent decree in San Francisco (Spitser 2006).[5] The nine years following the original consent decree significantly altered national educational goals and policies: the promise and prominence of integration faded. Studies in the 1980s had begun to examine the results of desegregation in terms of academic performance and found it lacking (Biegel 1988). The 1992 Orfield report followed suit concluding that while it met integration goals, it had "not succeeded in closing the [district-wide] achievement gap between whites and African Americans," (Emery 2002). Recall that while the consent decree required monitoring of student achievement, it neglected to establish specific targets, such as certain shares of students meeting a specific standard, or even goals, such as reducing the black-white achievement gap. While all schools integrated, not all schools successfully closed the achievement gap (Rojas 1996).

Orfield's report compared reconstituted schools to schools implementing less-stringent reform measures (Rojas 1996). Since the conclusions of the report led to wide-ranging decisions with broad consequences, discussion about the nature of this evidence follows an example (Figure 1) of the type of evidence presented.

[5] The report provided to Judge Orrick was titled "Progress Made, Challenges Remaining in San Francisco School Desegregation." Though this document is cited or summarized in many reports describing this era, access to this document is limited.

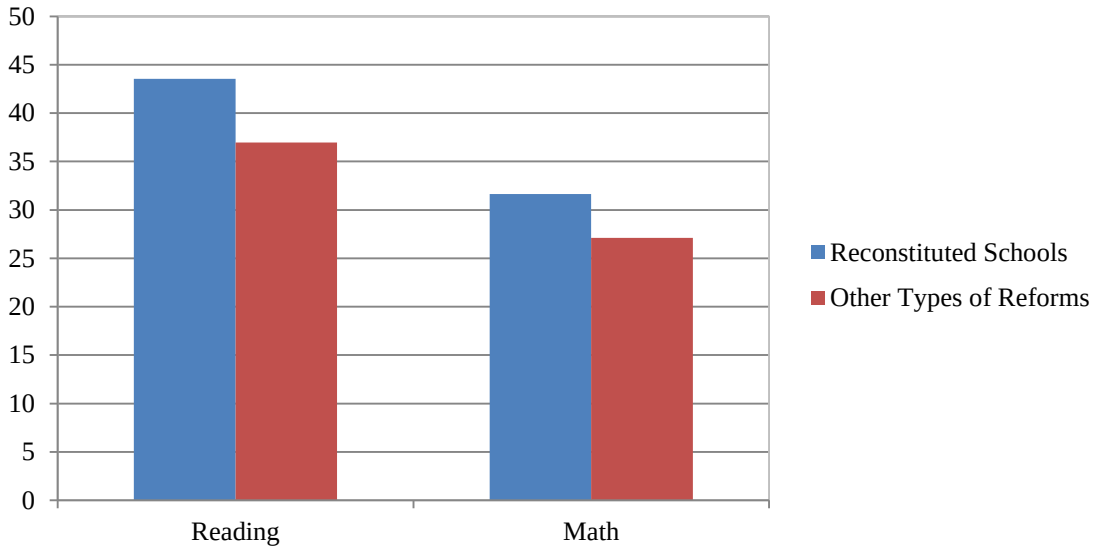

Figure 2.1 Mean NCE scores of African American and Latino students in San Francisco in 1992[6]

While apparently convincing at the time, demonstrated by its reproduction in various publications, this type of analysis includes several serious unaddressed flaws. Reconstitution occurred in 1984, almost a decade prior to these results. Results include a single year's data. Some reconstituted schools engaged in other types of reforms, including becoming a magnet or specialized curriculum school. Discussion of results lack any acknowledgement of the statistical techniques, so common in modern parlance, required to determine the influence of reconstitution on student academic performance. These facts alone provide reasonable doubt regarding reconstitution's ability to improve student achievement. While the report made over 70 recommendations, it reached one particularly important conclusion: since test scores at reconstituted schools surpassed those of non-reconstituted schools, the report recommended "that the district begin to reconstitute three schools each year," (Emery 2002, pg. 213).

[6] Re-created from Rojas, 1996, pg 103 Figure 9 (Rojas 1996).

Based on the evidence provided such a conclusion might appear unwarranted today. But the lack of 'real' evidence combined with the willingness of SFUSD to reconstitute more schools might originate from desperation. As standardized testing crossed the country in the nascent accountability era, many school districts found themselves in a precarious position. Schools clearly failed to meet standards and while their communities demanded improvements, few solutions to failing schools existed. Something, nearly anything, to address this problem possessed a broad appeal to communities across the country.

Regardless of its validity, SFUSD personnel viewed reconstitution as a means to improve student achievement (Rojas 1996). From this point forward, reconstitution's role in desegregation was largely ignored in favor of the narrative that it improved the academic achievement of students. The NAACP and SFUSD agreed to the modification of the consent decree allowing other schools to be reconstituted, though without specifying the mechanisms to choose them (Rojas 1996). The recommendations put forth by the 1992 Orfield report would soon encounter a powerful ally within the school district.

In August of 1992, SFUSD's Board chose Waldemar Rojas as their superintendent (Rojas 1996). According to the website Education and Democracy, "[i]n 1993, Rojas directed the development of the Comprehensive School Improvement Plan (CSIP). This was a process by which low performing schools were identified (primarily by test scores), offered resources and time, and then threatened with reconstitution if they did not improve student achievement" (Emery 2002). This implementation structure directly tied reconstitution to prior year academic performance. It also used reconstitution in two semi-divergent ways: as a potential reform, but also as a sanction. Teachers in underperforming schools could potentially lose their jobs. "Rojas argued in his 1996 PhD thesis, in numerous press conferences, and in all the official SFUSD publications that reconstitution

was responsible for raising San Francisco students' test scores generally and minority students' in particular," (Emery 2002). The immediate appeal of this measure quickly spread to other locations across the country and this structure, reconstitution as sanction, would find purchase in several states as they created their accountability systems during the 1990s.

Kingdon describes the criteria for a concept's survival into a short list of ideas as including 'technical feasibility' and 'value acceptance' (Kingdon 1984) . Elements of these criteria are included as justification for reconstitution in the dissertation of one of its strongest advocates, Dr. Rojas, the superintendent of SFUSD.

Kingdon defines technical feasibility as "attending to the feasibility of implementation, and specifying the actual mechanisms by which an idea would be brought into practical use,"(Kingdon 1984). In establishing reconstitution as a reform and sanction, the district needed to fully describe the school selection process and what reconstitution would look like. While the Orfield report advocated for reconstitution, it neglected to outline the mechanisms for identifying schools (Rojas 1996) . Though the report indicated that "reconstitution would include everything that was done for," schools reconstituted in 1984, those schools had engaged in numerous other reforms concurrently with reconstitution (Rojas 1996).

To determine which schools to reconstitute and how, the district developed "the Comprehensive School Improvement Program (CSIP)," (Rojas 1996, pg. 19). This plan included four components:

> (1) an identification and selection process of schools based on measurable indicators, (2) a one-year improvement period to develop and implement school change efforts by the school communities, (3) support to the sites by giving them special consideration in financial flexibility for their participation in school change initiative, and (4) participation in a review process to assess the rate of success of

> the reform efforts involving qualitative and quantitative indicators (Rojas 1996, pg. 65)

From these components the district developed additional technical details, including a description of the types of indicators used to identify troubled schools (Rojas 1996). At the end of each year, a host of criteria assessed whether a school made adequate progress to avoid reconstitution. Rojas (1994) includes lengthy descriptions of how the district identified schools for reconstitution and how the district informed communities and school personnel about the decision. These details included elements outside of the replacement of staff, including the campus-wide adoption of a set of philosophical tenets, flexibility in allocating staff time, staff development, and selection of the new curricular focus of the school (Rojas 1996). Working out these details contributed to the technical feasibility of implementing reconstitution in response to academic conditions at the school. After SFUSD implemented this version of school reconstitution, communities elsewhere intending to reconstitute their schools need only look to San Francisco, as the district developed, vetted, and implemented much of the detail work of reconstitution.[7]

Kingdon also mentions two sets of constraints to the implementation of new ideas: budget constraints and public acquiescence (Kingdon 1984). Rojas (1994) indicates that district policies addressed these constraints. The district prioritized budget requests from schools undergoing reconstitution over all other schools in the district (Rojas 1996). Additionally, "...schools were guaranteed no decline in categorical funding and had additional flexibility in the expenditure of the funds" (pg. 98). These guarantees effectively allocated resources to reconstituted schools; despite a drop in federal funding district-wide, schools engaged in the CSIP process maintained their funding levels (pg. 98).

---

[7] In fact, some scholars, in examining reconstitution in the 1990s in San Francisco took many of the implementation elements developed during this period (e.g., requiring schools to adopt a set of 'Philosophical Tenets'), and erroneously assumed these had been part of the original reconstitution endeavor in the early 1980s (Rozmus 1998)

According to Kingdon (1984), policies also require public acquiescence prior to implementation. Here, Rojas acknowledges some community concerns: "The introduction of the CSIP process generated bittersweet reactions throughout the District. On one hand, it opened the door to the reintroduction of reconstitution. On the other, it identified troubled schools..." (Rojas 1996) Here, Rojas hints at the community's positive view of reconstitution as implemented in 1983. Community resistance to the CSIP process lay not with expanding the number of reconstituted schools but with the public identification of poorly-performing schools.

The reforms implemented by SFUSD incorporated school reconstitution into a standard process, whereby the district offered schools, initially identified as failing to meet community standards, the opportunity to improve student outcomes first. If no improvement occurred, these schools would reconstitute.[8] This fundamentally altered the management and implementation of reconstitution from its original application in 1983. This revised version of reconstitution met the criteria Kingdon describes as necessary for this policy idea to coalesce into an accepted short list of acceptable alternatives.

Reconstitution would shortly spread like fire across the educational landscape of the US. Clearly, the need for a developed reform to improve failing schools possessed significant appeal. In our data-driven age, with more than a decade of high-stakes federal accountability policies, the lack of real evidence of the effectiveness of reconstitution in the early 1990s appears remarkable.

[8] This was an incredibly controversial transition. Students at one school designated for reconstitution marched on City Hall and enlisted the help of the Mayor of San Francisco to prevent their school from being reconstituted, all without success ibid..

The Diffusion of School Reconstitution

The 1990s saw the speedy diffusion of school reconstitution as a means of improving school-level student performance to many states and localities (Rozmus 1998). In Texas, reconstitution of schools occurred in Houston in 1994, Corpus Christi, Fort Worth and San Antonio in 1995, and El Paso in 1997 (Soledad 2006). By 1997, administrators reconstituted schools in Cleveland, Chicago, and Prince George's County, Maryland (Hendrie 1997). By the end of 1997, 23 states' accountability systems included reconstitution (Doherty and Abernathy 1998). Different locations experimented with reconstitution: applying it slightly differently, modifying how to identify schools requiring it, and incorporating reconstitution into state-level accountability systems. The spread of reconstitution moved faster than scholarly interest, so little careful examination and evaluation exists during this era.

But how does a reform effort develop broad appeal in communities as diverse as San Francisco, California and San Antonio, Texas? Kingdon notes that "...a complex combination of factors is generally responsible for the movement of a given item into agenda prominence," (Kingdon 1984). Reconstitution purportedly solved a critical public policy problem: failing schools. Principals, superintendents, and policy makers took claims of rapid success from San Francisco's reconstituted schools to heart, adopting a reform said to improve student outcomes. Results from recently created state-wide mandatory testing regimes and standardized national tests revealed that many schools performed below expectations. Reconstitution successfully traversed state and local borders because of an increasing sense of national crisis about education combined with a lack of solutions to failing schools.

One reason for the swift diffusion of accountability policies in general and reconstitution in particular, lay in the growing sense of the importance of education on the

national scene. Kingdon describes these "swings of national mood" (Kingdon 1984) as "the notion that a rather large number of people out in the county are thinking along certain common lines, that this national mood changes from one time to another in discernible ways, and that these changes in mood or climate have important impacts on policy agendas and policy outcomes,"(Kingdon 1984).

Gallup polls Americans on their view of the most important problem facing the country every year since 1949 (Gallup NA). Beginning in the mid-1990s, a growing share of respondents chose education as the most important problem facing the country (Gallup NA). Prior to 1991, only a very small share of respondents did so, between zero and two percent for any given year (Gallup NA). Between 1991 and 2000, the share of respondents who considered education the most important issue facing the country went from 4% to 10% (Gallup NA). By the turn of the century, one out of every ten Americans belived education was the most important problem facing the United States.

The first instance of incorporating reconstitution policies outside of California occurred in 1993, when the state of Maryland granted this power over failing schools to the State Board of Education (Cipollone 1998). Reconstitution consisted of several potential interventions including "changing one or more of the following: a school's administration, staff, organization, or instructional program," (Cipollone 1998, pg. 88). Already, key features of the governance structure invoking reconstitution and its implementation experienced significant changes from the San Francisco model.

The first recorded case of reconstitution applied to a school outside of San Francisco occurred in 1994 in Houston. Thomas J. Rusk Elementary School received "a negative report from a Texas Education Agency (TEA) accreditation team, complaints from parents and years of low scores on standardized tests," (Soledad 2006, pg. 117). A district-appointed principal implemented reconstitution reforms intended to improve the

school's performance and safety (Soledad 2006). Subsequent reports from U.S. Department of Education and TEA showed dramatic improvements in student academic performance (Emery 2002, Soledad 2006).

Reconstitution's watershed year occurred in 1995: efforts to reconstitute schools occurred in Boston, Maryland, and Texas (Rozmus 1998, Soledad 2006). Illinois state late law incorporated reconstitution as a sanction (Rozmus 1998). In Texas, schools were reconstituted in Fort Worth, Corpus Christi and San Antonio (Soledad 2006).

In Fort Worth, the superintendent identified schools for reconstitution, mandating 200 days of instruction as opposed to the minimally required 187 (Soledad 2006). The schools developed a comprehensive plan involving curriculum, student services, "staff development and defined approaches to student conduct," (Soledad 2006). Soledad records each of the schools' accountability ratings improved by the 1999-2000 school year. However, data from Texas' Academic Excellence Indicator System shows that all reconstituted schools academically were acceptable prior to the intervention (Soledad, 2006 #3; AEIS 1994-2011). Perhaps the driving force behind the superintendent's decision to reconstituted a reflection of low, but acceptable, performance on state-mandated tests rather than a desire to meet the threshold passing rate required in Texas.

Corpus Christi ISD reconstituted Wynn Seale Middle School in 1995 after being classified as low performing for the previous two years (Soledad 2006). Wynn Seale's curriculum changed into a fine arts neighborhood magnet school (Soledad 2006). Other structural changes occurred, including some block scheduling for core courses, an introductory arts program called Orbit that provided opportunities for sixth graders to explore creative options to pursue in the 7th and 8th grades, and curriculum changes that incorporated fine arts into the core subjects (Soledad 2006). The school went on to receive

accolades from various Texas educational organizations for student learning (Soledad 2006). These schools improved their accountability ratings (Soledad 2006).

In San Antonio, poor math scores and low morale led the school board to reconstitute Fox Tech High School (Soledad 2006). The district provided "a magnet coordinator and three instructional guides...to write curriculum and conduct staff development," (Soledad 2006, pg. 119). The new administration created smaller schools within the school with dedicated counselor, teacher, instructional guide and assistance principal staff within three learning academies: "a law and research academy, an applied technology academy, and a universal global academy," (Soledad 2006, pg. 120). During the 1997-1998 school year, the district acknowledged that Fox Tech was "the most-improved school," in the district (Soledad 2006, pg. 121).

That same year, Ysleta Independent School District (a school district near El Paso, TX) reconstituted Hacienda Heights Elementary after poor performance on statewide exams (Soledad 2006). This school adopted a literacy program developed in the early 1990s as a collaboration between the local university, the mayor, city council, and the chamber of commerce and continued to use it at least through 2006 (Soledad 2006). Using reading assessments early in the year, teachers and students agree on yearly objectives and students take periodic tests to determine their progress (Soledad 2006). Additionally, Hacienda Heights Elementary instituted dual language instruction (Soledad 2006). By 2000, the Texas Education Agency labeled the school exemplary and was designated a U.S. Department of Education National Blue Ribbon School (Soledad 2006).

Just eight miles away, across a set of railroad tracks, El Paso Independent School District's superintended ordered the reconstitution of Guillen Middle School and "required

all faculty and staff to reapply for their jobs," (Soledad 2006).[9] Incentives in the form of $5,000 stipends (nearly 17% of the average teacher salary of $29,590 that year at the school) were offered for teachers hired to work in the reconstituted school (Soledad 2006). Teachers at Guillen collaborated with University of Texas El Paso during the process of reconstitution and instituted several reforms such as creating longer class times during the day and mandating the use of uniforms (Soledad 2006). However, these efforts did not enable the school to meet their academic goals over the next four years (Soledad 2006).

The next year (1996) included proposals to reconstitute schools in Nashville, TN. Both Maryland and Wisconsin included reconstitution in their statewide plan for poor performing schools (Rozmus 1998). Superintendents and principals in Minneapolis, MN, Milwaukee, WI and Oakland, CA called for reconstituting failing schools (Rozmus 1998). Rozmus (1998) reported reconstitution of schools in Philadelphia, Atlanta, Houston, Memphis, New York City and Albany (Rozmus 1998). None of these efforts included documentation regarding implementation or studies determining the effects of school reconstitution on student performance.

In Texas that same year (1996) Anthony Trujillo, the superintendent of Ysleta Independent School District (the same one as in 1995) and the school principal decided to reconstitute Bel Air High School (Soledad 2006). Though classified as being academically acceptable using state standards, interviews with teachers revealed to the principal that "[t]here was an attitude that the kids...can only get to a certain level," (Soledad 2006). Unusually, Bel Air High School reconstituted in the middle of the school year (Retana 1997). After more than 50% of school staff left, achievement levels of the school improved (Soledad 2006).

[9] This type of reconstitution provided a principal with the option of keeping or removing whatever shares of staff the principal deemed necessary. This type of reconstitution places full authority of who remains and who leaves with the principal. State-mandated reconstitution beginning in 2003 follows this precedent.

In 1997, Chicago Public Schools (CPS) reconstituted seven high schools, which included firing most principals forcing teachers to reapply (Anagnostopoulos and Rutledge 2007). Being a large urban school district with a strong desire to reform poor performing schools and the eventual national prominence of the [Chicago Public Schools] CPS agenda, several studies examined the process of reform in Chicago (Anagnostopoulos and Rutledge 2007). Wong et. al. discussed the district-wide transition from a decentralized system of governance to a more centralized one starting in 1995 (Wong, Anagnostopoulos et al. 1999). Part of this re-centralization involved the launch of a local educational accountability system.

Clearly the policy of reconstitution, in the form developed in the early 1990s in San Francisco that involved probation and sanction possessed a broad appeal. However, these early implementers of reconstitution considered it simply one of a set of tools available to reform schools not meeting expectations. Reconstituted schools, by turns, changed into magnet programs, reclassified their grade levels, adopted new curriculum, and restructured to produce specific learning communities, just like schools reconstituted in 1983 but with arguably more flexibility. As evidence of many of these cases of reconstitution came from newspapers and from tangential references in the literature, the reconstituted schools discussed in this section represent a small fraction of the cases of implemented reconstitution. Any comprehensive analysis of reconstitution, its diffusion and effects, inherently contains selection bias: with visible cases weighted toward controversy and schools experiencing dramatic improvements. However, the examples above provide evidence that reconstitution often included other types of reforms and implemented for various reasons.

THE THIRD WINDOW

Whereas some states, like Maryland, adopted reconstitution in official law prior to its implementation, schools in Texas applied reconstitution first. In order for reconstitution to become state law, it needed a clear advocate to effectively craft and pass legislation. Just as Dr. Rojas advocated for the adoption of the CSIP process in San Francisco, a group of Texas legislators played a prominent role in bringing the option of reconstitution into legislatively required reconstitution. However, a brief statutory history of Texas and its school accountability system is in order.

Texas established its school accountability system in 1993 with the passage of Senate Bill 7. As noted in the previous section, a number of schools experimented with reconstitution though the state did not mandate it as a sanction. SB7 included other sets of statutes providing incentives and sanctions for schools and districts in Texas, based on a series of mandated tests in specific academic subjects (SB 7 1993). The statute mandates failing schools be subjected to a number of provisions:

> …if a campus has been a low-performing campus for a period of one year or more, appoint a board of managers composed of residents of the district to exercise the powers and duties of the board of trustees of the district in relation to the campus; or...if a campus has been a low-performing campus for a period of two years or more, order closure of the school program on the campus. (SB 7 1993)

SB1 in 1995 added the other possibility of "requiring the district to enter a contract for management of the campus by another person" (SB 1 1995). These limited options, closing a school, assigning a board of managers, and contracting out the school to another provider, hint at a loss of local control over education, a hallmark of Texas' public school system.

But advocates for the policy of reconstitution resided in the legislature. Beginning in the 1999 spring legislative session, Florence Shapiro, a state senator from outside Dallas, submitted SB 876:

> If a campus has been a low-performing campus for a period of two years or more, order closure of the previous school program on the campus or order that some or all of the educators employed on the previous school campus for at least the two preceding years be reassigned to other campuses or other appropriate positions in the previous school district. (SB 876 1999)

This language, while not invoking the word reconstitution, describes the removal of some or all of the teachers on that campus for the preceding two years (SB 876 1999). Clearly similar to the definition of reconstitution, the replacing of staff served as a sanction for schools classified as low-performing for several years.

Shapiro puts this proposal on the calendar on March 4th, relatively late in the semiannual Texas legislative session, which starts in January. On May 10th, just shortly before the end of that session, Eliot Shapleigh signs on as a co-author. Texas senator Eliot Shapleigh represents the area encompassing El Paso, a region where several well-regarded reconstitutions occurred. Though neither senator sat on the committee, the Senate education committee reviewed the bill:

> During "reconstitution" all or part of a campus' employees are required to resign and reapply for jobs with the district. In Texas, this process is only used in extreme cases. This bill would require the commissioner of education to establish a mandatory reconstitution process for chronic low-performing schools to attempt to address citizens' concerns over the plight of students who are unable to leave "low performing" public schools. (SB 876 Senate Bill Analysis 1999)

The committee recognized the intent of the proposed law required poor performing schools to reconstitute, though the word reconstitution never makes an appearance in the bill. However, the summary of the bill generates some general concern when reviewed by the House Research Organization (HRO). The HRO provides a summary of arguments that supporters and opponents of bills make. Supporters cited a study performed by the Texas Comptroller (SB 876 House Bill Analysis 1999). The 1999 Texas Performance Review, titled 'Challenging the Status Quo', included several recommendations related to education, one of which indicated "Clean the Slate at Substandard Schools," which cited

dramatic improvement in reconstituted San Antonio schools (Rylander 1999). Specifically, the report said: "Chronically low-performing schools should be required to reconstitute their staffs if they fail to improve after a probationary period. A school should be given a reasonable opportunity to improve student performance before its staff and administrators are reassigned," (Rylander 1999).

Opponents' views represented in the HRO indicated concern at granting such broad authority to the commissioner of education to reconstitute schools, and as such, "SB 876 would be a direct intrusion of state authority on local control and would directly affect the contract made between an educator and a local district." (SB 876 House Bill Analysis 1999). Though SB 876 passed the senate and house committees, it never made it to a house-wide vote.

During the next legislative session in 2001, Shapiro, now a member of the Senate education committee, files a similar proposal prior to the start of the session, in November of 2000. Shapleigh signed on as co-author in late January. This certainly afforded the opportunity for more legislative reflection on the proposed bill. The introduction of the bill included the following summary of its purpose:

> Currently, Texas does not have a program in place to reconstitute a low-performing school. There is no process to allow the commissioner of education to take measures to disband a school's faculty, including closing down the building and reopening the school with new staff, structure, and curriculum. As proposed, S.B. 26 directs a campus to be reconstituted if the school is classified as low-performing for two consecutive years. (SB 26 Senate Bill Analysis 2001)

The text of this bill was similar to that of the previous year, indicating that schools that fail for two years in a row either be closed or have "reassignment to other appropriate positions in the district of some or all of the educators employed for at least the two preceding years on the campus," (SB 26 2001).

This time around, the HRO did not provide an analysis of the bill and while it made it through the Senate and House education committees, it did not make it to consideration on the House floor.

Clearly, Shapiro pushed the reconstitution agenda. However, in each of these legislative sessions her efforts, while not meeting active resistance (as evidenced by the bill being voted positively out of the education committees in each instance) did not bear fruit in the passage of the bill. While the complex and obtuse political process in Texas may explain this, one reason for the lack of interest and passage might include the lack of articulation and urgency centered on a problem reconstitution solved. Additionally, while it provided the commissioner an alternative to closing failing schools, no one indicated the scale of need: how many schools might fall under this provision. Another problem leading to lack of legislative interest might include a lack of political support from other members of the legislature; each of these bills had a single co-author, potentially an indicator of a lack of enthusiasm for reconstitution more generally.

In the 2003 legislative session, Shapiro received the chair of the Senate education committee. Shapiro's strategy for pursuing the legislation she submitted in the past two legislative sessions changed substantially. Rather than submitting the legislation herself as author, her previous co-author Shapleigh submitted the legislation. Shapiro signed on as co-author. This bill also garnered three additional co-authors. Several substantive changes occurred in the proposed bill's language and in its introduction to the legislature that year. The revised language stated the following:

> If a campus has been a low-performing campus for a period of two years or more, the commissioner shall reconstitute the campus. In reconstituting the campus, a special campus intervention team shall be assembled for the purpose of deciding which educators may be retained at that campus. An educator who is not retained shall be assigned to another position in the district. (SB 618, 2003 #137)

The legislation now directly refers to reconstitution. In addition to reconstituting the school, though, it calls for a special campus intervention team, a feature not present in previous versions of this bill. The articulation of the bill in the introduction to the members of the Senate committee of this version also differed substantially from previously discussed efforts:

> Between 1994 and 2002, the number of low performing schools more than tripled, increasing from 54 to 149 campuses, including charter schools and alternative education programs. In 2001-2002, nearly 70,000 students attended public schools rated as low performing. Reconstituting schools that rank as low performing for two years or more is one of the five corrective actions required by the No Child Left Behind Act and is supported by the American Association of School Administrators. (SB 618 Senate Bill Analysis 2003)

This paragraph summary of the purpose of the bill ties the policy of reconstitution to the problem of failing schools, defining the scope of the problem both in the number of failing campuses and the number of students residing in failing campuses. This language significantly deviates from that previously used in bill explanations in 1999 and 2001, which lacked a clearly defined problem; clearly this same type of language (number of failing schools and the number of students attending them) if applied in previous years, could as adequately justify previous bills. However, the chances of passing this bill improved because the billl language aligned state accountability standards to the relatively new federal standards passed in No Child Left Behind.

Kingdon describes "spillovers" as when "the appearance of a window for one subject…increases the probability that a window will open for a similar subject," (Kingdon 1984, pg. 190). The election of 2000 brought the former governor of Texas, George W. Bush to the White House. One of the key early initiatives in his presidency was the signing into law of a revised version of the Elementary and Secondary Education Act (ESEA) that included provisions related to requiring standardized tests and accountability measures,

often referred to as the No Child Left Behind Act (NCLB) of 2002. States were required to fully develop an accountability system, and the next legislative session in Texas after the passage of NCLB occurred in 2003. This federal initiative which vaguely describes a reconstitution-like process as a sanction for the longest low-performing schools, helped to push along reconstitution measures within Texas for its state accountability system.

Florence Shapiro made a brief reply to comments from State Senator Shapleigh during the Texas Senate education committee meeting that addressed the bill that contained the reconstitution language on March 11, 2003. She indicated that "I do believe, Senator Shapleigh that No Child Left Behind addresses this very specifically. After a certain number of years there must be reconstitution so I think we're right in line with the federal government and the desires that they have to leave no child behind, truly," (Senate Committee Hearing at 18:29, transcription by the author).

In the four years since Texas state senator Shapiro put forth her initial reconstitution bill, the policy environment markedly changed. The first time she proposed this policy one of the chief objections included concerns that the state of Texas should not reach directly into the contracts between teachers and their districts. Such a provision overstepped the authority of the state. Just four years later, she hints that not merely state but federal policy could reasonably interfere with local control, citing this as a reason to bring state and federal policy into general alignment. This is a far cry from the primary concerns brought before the Texas legislature the first year she proposed this legislation.

Kingdon's multiple streams framework clearly explains the passage of this bill in Texas. This includes the articulation of a clear problem requiring a solution: the increasing number of failing schools. The policy solution met Kingdon's criteria for survival. Finally, a member of the Texas legislature doggedly pursued passage of a law mandating reconstitution over several legislative sessions. That legislative session, the bill passed into

law, requiring schools rated academically unacceptable for two years in a row to reconstitute. Since its passage in 2003 and prior to 2012, more than 135 schools reconstituted under this provision.

CONCLUSION

This chapter examined reconstitution's origins and describes how it became a feature of state-wide education policy in Texas. The first policy window created the first attempts at reconstitution in 1983 in the SFUSD. Reconstitution arose as a means of desegregating the lowest-performing schools in an often neglected neighborhood in San Francisco. The initial concept of reconstitution arose due to constraints on the court's ability to integrate across school districts and general opposition to integrating teachers. These first reconstituted schools also engaged in other reforms, such as creating magnet programs.

A review of the effectiveness of a host of school reforms in San Francisco after 1983 demonstrated that students attending formerly reconstituted schools demonstrated higher academic performance on standardized tests in 1992 than their peers at non-reconstituted schools. This review led to the second policy window, where the original concept of reconstitution dipped back into the 'policy primeval soup' and recombined with developing notions of school accountability into a clearly articulated probation and sanction regime. This process placed schools on notice for poor performance and required their reconstitution if they continued to underperform. This 'new' reconstitution policy spread rapidly across the country and many locales experimented with it.

The third policy window occurred in Texas in 2003 when the legislature passed a bill requiring reconstitution for schools classified as low-performing for two consecutive years. Texas schools experimented with reconstitution soon after San Francisco in the

early 1990s, though schools applied most of these early applications without the benefit of threatening a school prior to its implementation. Efforts, led by legislators, incorporating reconstitution into state law began as early as 1999. Several iterations of proposed laws during two legislative sessions led to its acceptance and passage into law in 2003. That policy window benefited from a clear articulation of the problem reconstitution was intended to solve, a persistent advocate who that year became co-chair of the senate education committee, and 'spillover' from recent passage of No Child Left Behind, itself advocated for by the former governor of Texas, George W. Bush.

One of the key reasons for examining the history of a policy stems from a desire to assess its effectiveness. Assumptions about the intent of a policy and reasons that the policy might successfully cause change inherently follow it as the policy diffuses, as with reconstitution. For example, the prerogative in the original implementation of reconstitution in San Francisco lay in integrating segregated schools; however, most post-1992 reconstitution cases lack mention of integration. Perhaps the initial success of reconstitution relied on integration of students and faculty. Thus, theories developed regarding the mechanisms of success for reconstitution should include those from the beginning of the intervention. Chapter three takes the context described in this chapter and focuses on the findings from the reconstitution literature, including a discussion of theories as to why school reconstitution might improve student academic achievement.

Chapter 3: Literature Review

Though requiring monitoring and reporting of student academic performance, the original cases of court-mandated reconstitution in 1983 provided no guarantee to a certain standard of academic excellence. Perhaps because of this, reconstitution lacked a fully developed or established theory as to how it might improve student academic achievement. Beginning in the early 1990s, reconstitution's link to improved academic performance occurred after its initial implementation. This left those studying reconstitution without a clear theoretical structure for how this measure could lead to potential improvements in student academic performance.[10] Without an established theoretical structure to test, researchers, time and again, engaged in theory building work or utilized theories from their own diverse fields to investigate reconstitution. Scholars describe the results of changes at schools on staff and, loosely defined, school climate, but only a few actually examined student performance measures; those that did so only examined aggregate school-level results. Little to no quantitative evidence exists linking reconstitution to student academic performance. The great irony of reconstitution remains that, despite the efforts of numerous scholars between 1983 and today, almost no quantitative work confirms or rejects the fundamental link between reconstitution and student achievement.

This chapter synthesizes previously developed theories from multiple fields related to school reconstitution, providing a number of reasons why reconstituting a school might improve student achievement. Evidence in the literature describes the effects of reconstitution on school staff and students, and discusses its limitations. The last section

[10] Established theory as per Brown v. Board suggested that integrating segregated schools would eliminate the inequitable treatment of students attending segregated schools, but this did not mean that it guaranteed equality of student outcomes.

of this chapter discusses the contributions of this work to the existing literature on school reconstitution.

THEORIES

Scholars identify three overarching theories to explain why reconstituting a school might improve student achievement.[11] Theories focus on reasons behind why a school might continuously fail, and each is aligned with a slightly nuanced perspective on the nature of learning with its origins in educational theories about the process of learning. The human capital hypothesis suggests that teachers at failing schools lack the skills needed to effectively teach their students. Replacing teachers who lack these skills with teachers that possess these skills should increase student performance. Principal-agent theory hypothesizes that inherent problems in the nature of teacher-student relationships prevent optimum learning. Replacing teachers changes this potentially skewed relationship, leading to improvements in student educational achievement. The racial balance hypothesis provides a conceptual link between student achievement and teacher integration, suggesting that inherent non-modifiable characteristics (specifically, race and ethnicity) play a critical role in the motivation of students to learn; replacing staff in such a manner that it creates racial balance should then improve student learning outcomes. Each of these theories possesses a natural affinity to a particular theory of learning.

Behaviorism and the Human Capital Hypothesis

Beginning in the early twentieth century, education scholars and psychologists developed a set of beliefs related to how individuals learn, with roots in the much earlier

[11] Some scholar's applied Argyris and Schon's theory of action (Koury 2000); however, others determined that this theoretical structure could not explain the reconstitution process, as "the dominant patterns of implementation we discovered…ran counter to the major premises of the policy," (Malen 2002). Furthermore, developed theories of action and theories of change often relied on fundamental human capital theory arguments (i.e., that replacing staff increased human capital resources at the school).

philosophical writings of John Locke and David Hume (Phillips 2014). Scholars refer to these collective sets of beliefs as behaviorism (Phillips 2014). The fundamental tenets of behaviorism stem from animal research related to the stimulus-response connection. Pavlov demonstrated that a consistently applied non-related stimulus (ringing a bell) could elicit the same natural response (salivating) as the true stimulus (food) engenders (Phillips 2014). Watson believed that like the shift from a natural reflex to a conditioned reflex seen with animals, human beings could learn more effectively, provided consistently applied stimuli and responses (Phillips 2014). Building on this work, E. L. Thorndike worked with cats, demonstrating that they repeated specific behaviors if they received some satisfaction as opposed to discomfort (Phillips 2014). B. F. Skinner took these animal models and developed the concept of a teaching machine: break complex information into more manageable learning pieces and provide rewards for individuals who mastered these smaller lessons (Phillips 2014). Building on Skinner's work, behaviorism purports that teachers hold the capacity to completely and effectively manage student learning, provided knowledgeable teachers create a proper system of challenge and reward (Phillips 2014). While a number of cogent and reasonable critiques arose out of this work beginning in the 1950s, leading to alternate or modified theories describing how individuals learn, the notion that teacher knowledge and behavior resides at the core of student learning remains a powerful force. From the behaviorism perspective, failing schools include a large share of teachers who are ineffective at their work and the fundamental cause of this lack of effectiveness stems from the teacher's inability to establish an effective system of teaching using appropriate stimulus and response within the classroom.

Behaviorism rationalizes that if teachers lack the knowledge and skills they need, students may not learn needed lessons. In order to correct their deficiencies, teachers may receive training to improve their teaching. This training, once applied in the classroom,

should improve student learning since the only inhibitor to student learning, according to behaviorism, resides in the lack of teacher knowledge and skills. Thus, increasing the human capital of teachers on a campus should lead to improvements in student learning.

Rather than spend significant time and effort determining and providing the types of training teachers need, school reconstitution allows a school to remove staff who lack the necessary human capital. Reconstitution thus increases the aggregate human capital on the campus available to students (Rice and Malen 2003). Malen (2002) summarizes the beliefs of individuals engaging in reconstitution, essentially reformers believed "that reconstitution would meet the immediate aim of creating more capable and committed faculty and staff " (Malen, Croninger et al. 2002).

Constructivism and the Agency Hypothesis

Somewhat counter to the notion of teaching machines, John Dewey espoused the belief that education should take into account a student's prior learning (Phillips 2014). Thus, the teacher must learn about their student and respond to the student's learning needs (D-Angel 2009). Constructivists believe that individual knowledge is removed from concrete facts that can (or must) be learned; learning should be centered on the individual and their construction of knowledge (Richardson 1997). Students construct their own knowledge while their teacher merely provides guidance. Knowledge does not exist outside a student's own construction. As learning cannot occur without the student's effort, students act as the primary agents of their own learning. Thus, teachers provide guidance to students, or help students manage their learning, but cannot mandate student learning as the behaviorism model suggests. "The teacher encourages [learning] in two ways: facilitating an environment in which students undergo a certain amount of cognitive

dissonance, and devising tasks that hopefully lead to a reorganization of existing cognitive maps," (Richardson 1997, pg. 5).

Thus, from a constructivist viewpoint, teachers do not merely present and test the knowledge students learn, but act as the principals of the work students (as agents) are asked to perform, which may or may not lead to student understanding depending on student goals and preferences. As in a traditional principal-agent problem, the goals of teachers and students may not align and teachers may not always be fully capable of determining if students engaged in the desired behavior, learning (Eisenhardt 1989). Teachers seek to maximize the learning of students through grade-based incentives, though students possess significantly more information about their willingness to perform the work required to achieve mastery.

When the goals between the principal and agent are perfectly or very closely aligned, the agent is more likely to perform as the principal desires (Eisenhardt 1989). Underperforming schools may experience more differentiation between student and teacher goals compared to well performing schools because students likely arrive at underperforming schools academically below grade level. Students arriving in the classroom below grade level may need to work harder to catch up to their current grade level. This additional effort likely represented a greater effort than the student applied in the previous year (evidenced by the student being below grade level). Thus, the teacher in this situation requires greater effort from the student in order to meet the teacher's goals, and the student (evidenced by the amount of effort applied in the previous year) may intend to keep the work level the same as that applied in the previous year.

In underperforming schools, the immediate consequences for poor student performance falls heavily on teachers, as they may lose their jobs. In the context of accountability systems, "policies that hold teachers accountable for student performance

intensify teachers' dependence on students to coproduce results," (Anagnostopoulos 2003). The principal-agent relationship of teachers and students implies that increased risks for student failure at underperforming schools as opposed to well performing schools, should drive teachers to pass as much risk as possible onto their students (Eisenhardt 1989).

Replacing staff allows creation of new types of relationships between teachers and their students without the weight of past assumptions about student academic capabilities. This rebooting of the teacher-student relationship positively changes the dynamic of the principal-agent relationship, one mechanism through which school reconstitution might improve student achievement.

Racial Balance Hypothesis

Also steeped in a constructivist view of educational learning, the racial balance hypothesis builds upon the legal theory behind Brown V. Board. While the 1954 court case established that separate education is inherently unequal, the racial balance hypothesis essentially extends the theory to describe how teacher integration acts to improve academic achievement (Spitser 2006). Teachers' internalized biases toward their students' academic potential at consistently low-performing schools prevents significant learning improvement; removing these teachers *en masse* modifies this detrimental teacher-student dynamic. There are two mechanisms through which this might work: 1) teachers with similar racial/ethnic characteristics consider their students more sympathetically and as such respond more capably to their needs, and 2) students see teachers who reflect their ethnic/racial characteristics as examples of success to which they may aspire (Spitser 2006).

These three theories comprise the essential arguments of reconstitution scholars. These theories naturally suffer from potential issues including measurement error. Perhaps

because of these limitations, many scholars focused on the sanction selection process, intermediate evidence of changes, and student achievement outcomes at reconstituted schools.

EVIDENCE OF THE EFFECTS OF SCHOOL RECONSTITUTION

Prior studies focused on the intermediate effects of reconstitution. These included changes in faculty and staff, perceived or observed student behaviors, and student perceptions of the reconstitution process. A small set of studies examined the relationship between reconstitution and student achievement using disaggregated data to examine school-level passing or sufficiency rates. Many studies attempted to establish whether and if the primary assumptions behind reconstitution's theories held: that replacing teachers would create changes at the school, which therefore might lead to changes in student performance. The following discussion describes scholar research on the reconstitution selection process, the changes on campus following reconstitution, and evidence of the effects of reconstitution on student achievement.

Selection of Reconstitution

Prior studies established that the selection process for reconstitution varied, including the primary decision maker. Discussion in the previous chapter mentioned these varied reasons and processes. Beginning as a mandated action through a court-based consent decree, superintendents often led the charge to reconstitute schools in the 1990s. Often schools targeted for reconstitution experienced years of poor performance prior to the initiative, which triggered community, school, district, and state concern. After the original cases of reconstitution, two selection processes determined which schools to reconstitute: either a sanctioning process within the school district with soft or hard triggers, or a demand by superintendents or principals for reconstitution.

District officials in the early 1990s placed schools in San Francisco on a probation list and forced non-improving schools to reconstitute (Rojas 1996). Schools in Chicago went through a similar process to those in San Francisco after the entire district underwent a change in governance process in 1995 (Hess 2003). That year in Chicago, a new "law enhanced the power of the [school] Board of Trustees and the CEO to identify low-performing schools and place them on remediation, probation or reconstitution," (Wong, Anagnostopoulos et al. 1999). The following year, the Board placed 109 low performing schools, 20% of all schools in the district, on academic probation (Wong, Anagnostopoulos et al. 1999, pg. 13). The next year, they demanded the reconstitution of seven high schools (Wong, Anagnostopoulos et al. 1999).

In some districts, the superintendent determined which schools would reconstitute. Respective superintendents mandated the reconstitution of schools in Maryland and Texas (Malen, Croninger et al. 2002, Soledad 2006). Occasionally, superintendent efforts staved off even more intrusive efforts by the state to close or further censure the school (Malen, Croninger et al. 2002). Scholars noted that though the superintendent ordered reconstitution this simplifies an often complex process involving discussions and agreement with the district board, teachers unions, and state authorities (Rice and Malen 2003). Occasionally principals demanded and received authority to reconstitute their school (Retana 1997, Soledad 2006).

Beginning in Texas starting in 2003, schools classified as academically unacceptable for two years in a row automatically faced reconstitution. This mandatory implementation rejects any principal, school board, superintendent, or state agency input related to the need to reconstitute. This type of governance structure removes choice, but

provides political cover for those who implement reconstitution: the school failed and must be reconstituted.[12]

Each selection method includes benefits and disadvantages. A loose governance structure where school boards decide on the how and where of reconstituting schools may provide opportunities for community input and reflection. Board members may experience political blowback from their decision, particularly if purported improvements never materialize. A strict governance structure where school performance triggers mandated reconstitution provides political cover for the action, but schools close to passing the threshold may not need such a wide-ranging intervention to improve. Spitser (2006) notes that because test scores often trigger reconstitution, they should accurately measure teacher effectiveness, which they rarely do. Cases where principals or superintendents directly order reconstitution places credit for success and blame for failure squarely in their court: failure of perceived improvements after reconstitution may cost them their jobs.

In part, differences in how to choose schools for reconstitution raise the specter of selection bias: schools with very capable or politically connected principals who encountered faculty resistance to their agenda could turn to a supportive board or superintendent to bring about reconstitution (Retana 1997, Soledad 2006). Superintendents receiving flak from their school board about a particular school could demand its reconstitution, if merely to provide evidence of effort to board members (Rice and Malen 2003). If non-academic processes select schools for reconstitution then the process and results of the intervention might also be guided by non-academic goals. Invoking a universal process involving standardized test score criteria prevents politically connected

[12] Federal reconstitution regulations include more options for schools at each stage of failure, which includes potentially reconstituting a school. These options make identification of reconstituted schools problematic. Texas' mandatory reconstitution policy provides no options and easy identification.

faculty and principals from hindering its trigger and treats all schools within the framework equally.

Evidence of Change

Scholars noted evidence of changes at reconstituted schools, some providing information about the school prior to and after reconstitution. All changes, save reconstitution-mandated staff turnover, likely occurred at other non-reconstituted schools across the country. Schools engaged in other types of changes and reforms concurrently to reconstitution. Scholars noted financial changes, changes in actual or perceived student behaviors, changes in school leadership, and teacher changes at reconstituted schools.

Financial Changes

The primary focus of this work spans the era between the first cases of reconstitution and the federal School Improvement Grant (SIG) era of reconstitution, which both provided generous fiscal resources to reconstituting campuses. Because of the relatively large set of fiscal resources available, these examples diverge from the primary subject of this dissertation. Superintendents and principals in these outlying cases applied the additional fiscal resources to search for and hire teaching candidates from across the country (Rojas 1996).

Schools reconstituted in Texas after 2003 did not receive significant changes in fiscal resources. In other cases across the country, schools receiving additional funds relied on grants to provide them. These grants included restrictions on how funds could be spent: these restrictions "undermined the school's freedom to align these resources with their own priorities," (Rice and Malen 2003, pg. 14) After reconstitution, additional fiscal resources provided by the district to reconstituted schools ended; the entire burden of replacing them fell to school staff (Rice and Malen 2003). Where the district provided additional

resources, some principals fought to maintain them in future years (Wong, Anagnostopoulos et al. 1999).

Student Behaviors and Identification

Prior to reconstitution, scholars note high rates of poor student behavior, such as roaming the halls, fighting, drug use, gang affiliation, and/or guns brought by students onto the campus (Retana 1997, Rozmus 1998, Hamilton, Heilig et al. 2014). A key aspect of the work of principals prior to reconstituting a school included assessing school climate and student behavior and developing an approach to deal with these issues (Koury 2000). Reconstituted schools with student discipline problems engaged in a host of initiatives to improve student and staff safety, including both physical security and engaging in developed conflict resolution processes.

Physical security sometimes involved installation of metal detectors and use of clear backpacks (Hamilton, Heilig et al. 2014). More often, physical security involved locking students out of classrooms during class, 'sweeping' the halls, and having staff monitor exits (Retana 1997, Soledad 2006). Some principals emphasized conflict management skills among students in a framework of emphasizing respect and dignity (Koury 2000). Students who experienced reconstitution sometimes felt their school was less safe than before (Rozmus 1998). However, scholars visiting reconstituted schools often noted that schools felt safe and secure (Retana 1997, Koury 2000, Malen, Croninger et al. 2002, Hess 2003, Rice and Malen 2003, Hamilton, Heilig et al. 2014).

School Leadership

Sometimes school boards or state agencies required superintendents to hire new principals as part of the reconstitution process, though often these principals left shortly after reconstituting the school (Hess 2003, Soledad 2006, Hamilton, Heilig et al. 2014).

Reconstituted schools sometimes exhibited evidence of high principal turnover rates at schools long before reconstitution, which occasionally included principals turning over multiple times within a year (Goldstein, Kelemen et al. 1998, Malen, Croninger et al. 2002, Soledad 2006, Hamilton, Heilig et al. 2014). In some cases, their new status made it difficult for principals to influence school decisions without resistance, as they had hired few of the staff (Hamilton, Heilig et al. 2014).

Teacher Changes

The primary mechanism of change at reconstituted schools involves replacing teachers and staff. The shares of teachers new to campus varied by location and circumstances, with roughly two-thirds being new to the campus in some instances (Wong, Anagnostopoulos et al. 1999, Malen, Croninger et al. 2002, Rice and Malen 2003, Rice and Croninger 2005) but lower in others (Hess 2003). Certainly, reconstitution removed a small number of truly 'dreadful teachers' to the benefit of students (Retana 1997, Hess 2003, Soledad 2006). Though scholars questioned if removing a smaller number of the poorest teachers demanded reconstituting the entire staff of a school (Malen, Croninger et al. 2002, Rice and Croninger 2005). However, some teachers indicated that leaving faculty applied older pedagogical practices which hurt student academic performance (Hamilton, Heilig et al. 2014). In one case, a superintendent viewed most teachers at one school as less dedicated to student learning than required (Malen, Croninger et al. 2002).

Principals required significant time to hire new staff, taking them away from other important activities (Rice and Malen 2003). Many principals at reconstituted schools faced challenges in finding qualified teaching candidates to fill vacancies (Rozmus 1998, Wong, Anagnostopoulos et al. 1999, Hess 2003, Rice and Croninger 2005, Spitser 2006, Hamilton, Heilig et al. 2014). Some scholars noted that the lack of qualified teachers led

to the school's failure, and replacing poorly qualified teachers with similarly qualified teachers might not improve student learning (Spitser 2006). Principals also experienced difficulty in effectively determining which of the existing teachers they should keep and which they should fire (Retana 1997, Hess 2003). While some schools hired additional positions during reconstitution, available funds came from grant funding or district programs and allow for only specific purposes, not necessarily those that campuses needed most (Rice and Malen 2003).

Newly hired teachers needed more time to build their skills and possessed less human capital than their departing counterparts (Rozmus 1998, Malen, Croninger et al. 2002, Rice and Malen 2003, Rice and Croninger 2005, Hamilton, Heilig et al. 2014). Principals at reconstituted schools required teachers to focus on improving student test performance leading to wrote and simplistic work, leading some capable teachers to leave (Wong, Anagnostopoulos et al. 1999, Spitser 2006). Spitzer (2006) posits a downward spiral of continuously reduced human capital as reconstitution drives out experienced, capable teachers and replaces them with less experience and less capable teachers over successive years. After reconstitution, those experienced teachers who initially stayed left due to district promised but unrealized supports (Malen, Croninger et al. 2002).

Though an often stated goal of reconstitution included the creation of an interconnected faculty, large shares of new staff made it difficult to create collegial networks of teachers and staff within the reconstituted schools (Retana 1997, Rice and Malen 2003, Rice and Croninger 2005). However, some teachers viewed the new teachers as enthusiastic and more capable of collaborating with other teachers than those who left (Hamilton, Heilig et al. 2014). Teachers that did keep their positions sometimes experienced "grief-induced guilt" about not being let go (Hess 2003, pg. 8).

After reconstitution, the gender and racial/ethnic characteristics of teachers aligned more closely to those of students than before at some schools (Malen, Croninger et al. 2002, Rice and Croninger 2005). Principals at some reconstituted schools hired staff from diverse backgrounds in an effort to nurture cultural understanding while emphasizing teacher responsibility (Rozmus 1998, Koury 2000). Retana (1997) indicated one school's reconstitution dealt a fatal blow to racist attitudes and actions by faculty members.

When examining traditional measures of teacher effectiveness, including certification, experience, and reputation, newly hired teachers after reconstitution reduced these characteristics within a school (Malen, Croninger et al. 2002). Malen (2002) came to the conclusion that the chaos engendered teachers to "reestablish basic operating procedures and revert to prior, familiar practices," which limited the amount of change experienced on a campus (Malen, Croninger et al. 2002, pg 124). Continued staff turnover, even after reconstitution, "made it difficult for principals to recover, let alone improve the stock of human capital at their schools," (Rice and Croninger 2005, pg. 88)

Though a goal of reconstitution often included increasing the availability and quality of training for teachers, relatively little evidence of such professional development exists. Retana (1997) found evidence for multiple types of professional development through the establishment of peer networks and utilizations of standardized course-level instructional guides. For schools that provided additional training for teachers the amount of time devoted to training added to an already full schedule of work, occurring throughout the year both before school and after school and on the weekends; principals required new teachers to engage in even more training (Rice and Malen 2003, Rice and Croninger 2005).

Student Academic Performance

Schools experiencing reconstitution rarely exhibited notable improvements in student academic performance, and their failure to improve sometimes continued for years following reconstitution (Hamilton, Heilig et al. 2014). In the early stages of San Francisco's reconstitution efforts in the 1990s, elementary schools experienced moderate gains in reading, but none in math, with other types of schools demonstrated mediocre gains (Rozmus 1998). Students that reconstitution intended to help in San Francisco (African Americans, Filipino, and Latino students) "failed to show growth" in reading and math test performance (Rozmus 1998).

In Chicago, seven reconstituted schools improved student reading performance, compared to a national norm, though this increase did not match that observed in other, non-reconstituted schools and even with improvements saw only small gains (Hess 2003). However, the median percentile in both reading and math at reconstituted schools experienced gains at the same rate as the rest of Chicago schools (Hess 2003). Results by quartile demonstrate a significant decrease (58.5%) in the number of students in the lowest quartile (Hess 2003). However, these improvements in academic performance "more reflect the better preparation of entering freshman than improvement in learning once students enrolled in" the reconstituted schools (Hess 2003, pg. 27). Wong (1999) examined the gap in passing the reading exam between reconstituted and non-probationary schools in Chicago, finding that the gap reduced from roughly a 1 to 8 ratio to a 1 to 5 ratio, with similar improvements in math scores. However, non-reconstituted schools placed on probation also improved at similar rates (Wong, Anagnostopoulos et al. 1999).

Malen (2002) indicated that schools followed their own unique and erratic patterns of student academic performance. After the first year, two schools experienced satisfactory performance on state tests, while the other four witnessed declines in performance (Malen,

Croninger et al. 2002). While some of the schools improved aggregate student achievement, such improvements proved unreliable in later years (Malen, Croninger et al. 2002). Retana (1997) indicated that student groups attending the school during the reconstitution year improved school-level passing rates in reading (for all groups of students), and in math (except for African American students), though student writing passing rates all declined (Retana 1997).

Thus, previous literature provides cogent information on how schools were selected for reconstitution, provides evidence of changes on reconstituted campuses, and hints at the effects of reconstitution on school-level student achievement. While these studies provide important context, they suffer from several limitations.

LIMITATIONS OF THE LITERATURE

Unfortunately, previous studies insufficiently establish evidence of change at schools, whether student achievement improved, or whether reconstitution caused any perceived improvements. Studies suffer from two types of limitations: those related to the nature of studying small numbers of schools experiencing reconstitution, and limitations related to the measurement of student performance. The first of these limitations prevent any determination of the generalizability of the results of any analysis. The latter type of limitations prevent drawing any conclusions as to whether reconstitution improved student outcomes at all.

Scholars examining school reconstitution focus on usually one, though sometimes up to eight, reconstituted schools within a generally small geographic frame. By only examining reconstituted schools, scholars neglect the possibility that changes are repeated in schools not undergoing reconstitution. From this perspective, attributing changes seen within schools to reconstitution may be illusory. Additionally, the small number of schools

examined by any particular scholar make conclusions about student improvements difficult to ascertain: student improvements in a single school may not be readily replicated in others. If more than a single school was studied, they all existed in a very similar geographic area. Conclusions about the effect of reconstitution in urban Chicago may not be relevant in suburban Springfield.

A few studies, as mentioned, use aggregate school-level results of student standardized test performance to determine whether school-level results improved after reconstitution (Wong, Anagnostopoulos et al. 1999, Hess 2003). None of these studies examined changes in individual student test performance, and so never attempt to measure the effect of reconstitution on those supposedly the subject of its treatment. School-level aggregated test performance might appear to indicate the success or failure of reconstitution, after all, the point of the intervention is prevent the school from failing. This does not truly address the counter-factual: i.e., what would have happened at these schools absent the intervention. At some schools, for example middle schools, a third of students to that campus are new that year. Thus, during the year following reconstitution, students attending a middle school for the first time have no experience with the middle school prior to reconstitution and so, it might be argued, have not actually experienced the treatment. Incoming students might be better or less prepared than their previous cohorts which is not accounted for in the analysis.

Unfortunately, the literature provides no established statistical procedures or estimates of the effect of reconstitution from which future scholars may anchor their research. In part, this deficiency is related to the small number of schools examined and the generally qualitative methods applied within those schools seeking evidence of change.

CONTRIBUTIONS TO EXISTING LITERATURE

This dissertation contributes to the field by including a thorough historical construction of local and state reconstitution up to 2003. Prior to this dissertation, scholars hinted at the long history of the application of reconstitution, perhaps even referenced the progenitor cases in San Francisco, but the reasons for its general appeal and potential explanation for its diffusion remained absent in this field prior to this work. This absence of historical context led many scholars to, time and again, develop or re-develop theories as to why reconstitution might improve student outcomes. This dissertation recognizes that these theories all possess fundamentally similar characteristics and may be reasonably tied to theories related to how students learn and how teachers interact with their students.

Despite over 30 years of reconstitution, basic answers to fundamental questions regarding changes in reconstituted schools remain: what share of staff leave a reconstituted school; when do they do they leave; what types of staff stay compared to the types who leave; do aggregate school-level results see improvements after reconstitution? While previous scholars established high teacher turnover at reconstituted schools, these cases include different governance selection processes leading to a possibility that schools may have been selected either purposefully or accidentally by researchers because of their strident efforts to replace staff. This work, by relying on 137 reconstituted schools across Texas between the years 2003 and 2011 allows for a more accurate portrait of reconstitution.[13]

Using individual, standardized state-wide student achievement data in mathematics and reading allows for a clearer description of any influence reconstitution might have on student learning. Previous work relied on measures that included the share of passing students from specific groups, or single-school measures describing success or failure at

[13] This one study includes more reconstituted schools than all other previous cases combined.

meeting performance targets. While these provide helpful context, they neglect whether school reconstitution actually improves student achievement, which only individual data can provide.

Being the first effort to examine large numbers of reconstituted schools within the same accountability regime using individual student data, this work provides the basis for all future attempts at measuring the effect of reconstitution on student achievement. The context and results provided in this dissertation allow future scholars the opportunity to critique and enhance the methods applied here.

Chapter 4: Methods

The first three chapters focus on the diffusion of reconstitution and theory development, answering several research questions from the introductory chapter. This chapter describes the quantitative methods intended to measure the effect of school reconstitution on student achievement in Texas. The first part discusses the research questions and outlines a number of potential hypotheses. The data section includes a discussion of the state accountability system in Texas between 2003 and 2011, during which all students completed the Texas Assessment of Knowledge and Skills (TAKS) tests. The data section describes all dependent and independent variables, spending additional time on constructed variables. The statistical analysis section describes the methods employed to answer the research questions. Methods include two regression discontinuity designs at the school level, along with several student-level fixed effects models with individual and time fixed effects. The last section describes data and methodological limitations of this study.

RESEARCH QUESTIONS AND HYPOTHESES

This study seeks to determine whether reconstitution achieves its goal of improving student performance on state mandated tests. Assuming that reconstitution plays a role in student performance, this study attempts to attribute where possible the mechanisms through which reconstitution works, be it changes in school-level human capital or teacher-student racial balance. The analysis presented informs the policy debate related to reconstituting schools. The guiding research question of this dissertation asks: What is the effect of the policy of reconstitution on student achievement in Texas? This overarching question consists of three separate research questions:

- How is reconstitution applied in Texas? (RQ1)

- Does reconstitution help failing schools improve their minimum passing score? (RQ2)
- What is the effect of reconstitution on student performance on standardized mathematics and reading tests? (RQ3)

Each of these research questions consist of more detailed efforts and, in some cases, lead to specific hypotheses related to the effects of this policy.

RQ1: How is reconstitution applied in Texas?

As demonstrated in the previous two chapters, the implementation of reconstitution varies depending on the location and governance structure of the state and school district. In Texas, principals assigned to a school engaging in reconstitution possess the broad authority to replace as many or as few staff members as they see fit. Since Texas reconstitution policy permits this flexibility and places no requirements related to the share of new staff, questions remain about whether and to what extent reconstituted schools replaced staff, whether staff after reconstitution possessed different characteristics than staff prior to reconstitution, and the variability of these changes across reconstituted schools.

RQ2: Does reconstitution help schools improve their minimum passing score?

While the focus of this work considers the effects of reconstitution on student achievement more broadly, the intent of the policy arguably takes failing schools and improves the measure used to assess their passing or failing status. From this perspective, the single failure measure used to determine passing or failing for schools (their minimum passing score, explained in detail below) should, after a school reconstitutes, improve.

RQ3: What is the effect of reconstitution on student performance on mathematics and reading tests?

Due to the structure of assigning failure or passing to schools in Texas, school failure or success is not necessarily perfectly correlated with student failure or success. Regardless of the answer to RQ2, students may (or may not) receive a benefit for attending a reconstituted school. The first corollary research question simply asks if there is any evidence of improvement in math and reading performance for students who attend a reconstituted school? (RQ3.1)

Since schools possess limited resources and reconstitution in Texas does not guarantee additional funding, staff within schools may choose to focus on specific populations of students in an effort to improve the school's chance of passing in the next year. The most obvious path of intervention might be to focus efforts on those students whose group failed in the previous year. Students in these groups, with the lowest passing rate at the school, arguably should see noticeable improvements. Do students who belonged to the group with the lowest test performance see improvements? (RQ3.2)

Other than focusing on the student group with the lowest passing rate, a school might distribute their resources to improve instruction in the school's worst-performing subject from the previous year. These efforts might be broadly based across the school or focused specifically on students who failed the subject in the previous year. Regardless, these efforts should lead to improvements for students who failed the test in the previous year. Do students who failed the subject-specific test whose failure triggered reconstitution see a benefit from reconstitution? (RQ3.3)

Reconstitution might generally provide more benefit for specific groups of students. As theorized in the previous chapter, schools reconstituted to provide racial balance might see improvement specifically because the staff now reflects the diversity of the student body. Hence, does reconstitution provide greater benefit for specific groups of students? (RQ3.4)

Students experience reconstitution as one element of the arc of their public education; students attend several schools, one of which is reconstituted either immediately before or during their tenure. The next set of research question corollaries considers the cohort nature of the experience of students at reconstituted schools. If it is the experience of reconstitution which leads to improved student achievement, then students who enter the school for the first time after it reconstitutes would not receive as much benefit as those who were previously enrolled at the school. Does reconstitution improve performance for students who experienced reconstitution (i.e., attended the school prior to and immediately after reconstitution)? (RQ3.5)

These questions hint at mechanisms through which the school selectively provided limited resources; they still consider reconstitution as a 'black box,' intervention. Identified effects of reconstitution on academic performance, while informative, need additional investigation examining the association of any potential types of campus-level change responsible for improvements. This dissertation determines which measurable factors identified in the literature attenuate student test performance due to reconstitution: specifically the share of particular types of staff new to that campus that year and their metrics of human capital; as well as the gender, racial, and ethnic similarity between staff and their students.

UTILIZED DATA

Data for this dissertation include records for over 6 million test-taking students in over 8,000 schools in Texas from 2003 through 2011, a total of over 26 million person-year records, accessed through the Texas Education Research Center (ERC) and combined with publicly-available campus-level testing results to determine which campuses Texas' state accountability systems required to reconstitute. This section provides an overview of

the data available to answer the research questions. It begins by providing a brief overview of Texas' state accountability system, providing context for the process used to trigger reconstitution. Description of the data and its structure, including descriptions of constructed variables follows.

Texas' State Accountability System

Texas state accountability ratings trigger reconstitution. Schools that fail (meaning they are classified as Academically Unacceptable (AU)) for two years in a row take a year to plan their reconstitution and then undergo mandated reconstitution the following summer. This section describes the accountability system determining a campus' classification, which then may trigger reconstitution.

Texas established its own accountability system in the early 1990s, nearly a decade before the passage of No Child Left Behind (NCLB) (Soledad 2006). The state modified accountability tests after passage of NCLB, adopting the Texas Assessment of Knowledge and Skills (TAKS) test in 2003. These new TAKS tests include additional content and more stringent standards than the previous Texas Assessment of Academic Skills (TAAS) test.[14] The Texas state accountability system after 2002 included a number of unique features not included in the federal accountability system. Rather than focus solely on mathematics and reading tests, the state accountability system also considers tests on writing, social studies and science. Depending on the year and the type of school, additional requirements for academic sufficiency include completion and dropout metrics. Campus passing rates on each of these tests for the entire school and for specific

[14] In 2011, Texas adopted the State of Texas Assessments of Academic Readiness (STAAR) test. Just as TAAS and TAKS tests lack statistical equivalency, so to do the TAKS and STAAR tests.

populations of students must meet an established standard; Academically Unacceptable (AU) schools fail to meet the minimum passing rate on any of these measures.[15]

The minimum requirements to avoid failure differ for each test and sometimes changed from year to year. For example, in 2005 schools needed to ensure that at least 35% of their students passed the math TAKS and 50% of their students needed to pass the reading TAKS test in order to be classified as Academically Acceptable. In 2006, schools needed a threshold of at least 40% of their students passing the math and 60% passing the reading TAKS test. The long-term goal of increasing the required passing rates followed federal initiatives to increase the minimum passing rate for all tests to 70% by 2015. Figure 4.1 provides the required campus-level passing rate in each year for each subject.

[15] Beyond the standards required to meet Academically Acceptable (AA), schools could be classified as being Recognized (R) or Exemplary (E). Each of the thresholds from AU to AA, from AA to R, and from R to E, included specific passing rate requirements and standards that needed to be met in order for schools to reach that evaluation level. As school failure provides the focus of this work details give information related to the 'passing' threshold from AU to AA; other thresholds are only discussed cursorily.

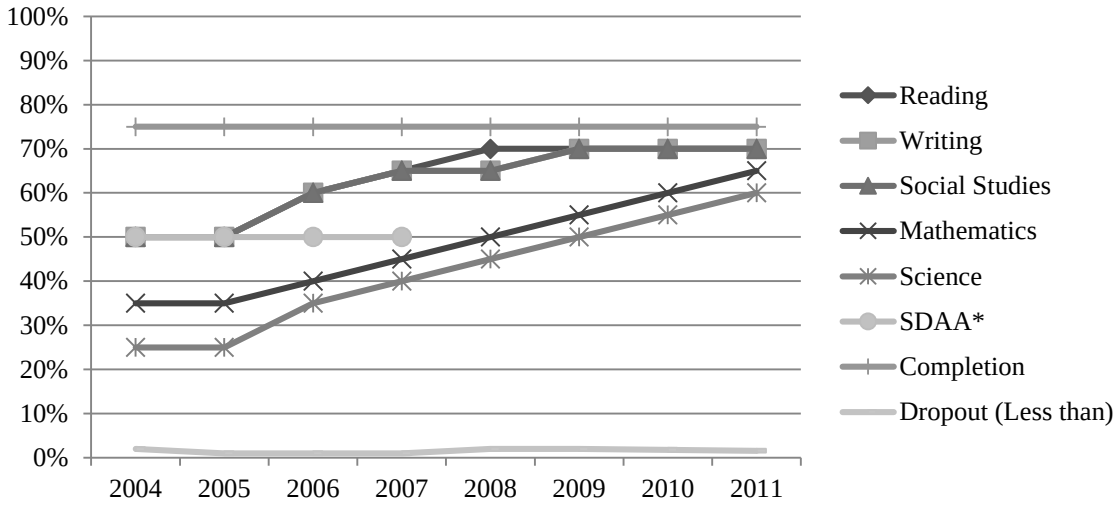

Figure 4.1 Required TAKS Passing Rates, by Subject and Year[16]

Not all students take all test subjects in each year. Students younger than 3rd grade take no TAKS tests, nor do 12th graders who passed the 11th grade exams. Table 1 describes which grades tested which students in which subjects. Except for reading/English Language Arts (ELA) and mathematics, students only intermittently take tests in other subjects. So an elementary school failing to meet the passing threshold in writing (or science) tests an entirely different group of students in that subject the following year. Thus, improvements in school-level test results may not necessarily represent educational progress at a school. This intermittent testing structure means that it takes as many as three years to effectively measure individual-level student improvement in science, writing, and social studies. Students failing to meet science standards in 5th grade test again in science at their middle school in the 8th grade. The most consistently proctored tests, math and reading, provide the clearest information on individual academic progress. The span of

[16] Figures 4.2 and 4.3 use information compiled from Texas Education Agency Accountability Manuals from 2004-2011. Note that SDAA were specialized tests for students classified as special education and requiring specific testing modifications. SDAA requirements ceased starting with the TAKS Accommodated tests in 2008.

years where students took the TAKS test cover 2003 through 2011, so students in the 3rd grade in 2003 experienced the entire arc of TAKS standards through to their junior year of high school.

Grade	3	4	5	6	7	8	9	10	11
Reading/ELA	x	x	x	x	x	x	x	x	x
Writing		x			x				
Mathematics	x	x	x	x	x	x	x	x	x
Social Studies						x		x	x
Science			x			x[17]		x	x

Table 4.1 Test Subjects, by Grade Tested

Between 2004 and 2007, some schools needed to meet requirements on six standards (math, reading, writing, social studies, science, completion and dropout thresholds) for five specific populations of students (all students at the school and low income, African American, Hispanic, and white students), for a potential total of 30 measures (measures being the combination of tests or standards for each of the specific populations of students). If a school failed to meet a single measure out of these 30, but passed all other measures, it would fail and be classified as AU. Theoretically, this meant that a school's minimum performance across all measures determined whether or not a school met state standards.[18]

In practice, a series of automatically applied provisions could shield the true minimum score from consideration for the purposes of classifying a school's academic status. The state progressively applied each of these shielding provisions for each campus, and, for certain provisions, solely to campuses whose minimum measure after the

[17] Students in the 8th grade were required to take the science test beginning in 2008.

[18] Federal provisions related to Adequate Yearly Progress (AYP) contain their own structure and implementation quirks.

application of all previous provisions still resided below the state standard.[19] Applying all standard provisions (described in detail below) meant that between 2004 and 2011, minimum scores below the passing threshold were changed to passing 2,456 times at 1,626 campuses.[20] Some campuses benefited more than once from these provisions. For 12 campuses in Texas, provisions moved the campus from unacceptable to acceptable status for five of the nine years examined.

Provisions include Required Improvement (RI), the Texas Projection Measure (TPM), and the Exceptions provision. The first such provision, Required Improvement, meant that if a school's calculated improvement on a measure from last year to this year, if repeated next year, would bring the campus to Academically Acceptable status, then that measure 'met' the standard. Required Improvement counted for any test and multiple tests at a school for any population of students, even to completion and dropout standards.[21] Between 1.8% and 4.9% of campuses shielded their minimum score from consideration using RI, depending on the year. Just over two-fifths of these changes to the minimum altered the specific population of students at the minimum, but not the test; for example, a school's minimum measure went from being math for economically disadvantaged students to being math for African American students. Another 31% altered the test, but not the specific population of students at the minimum; for example, a school's minimum score

[19] Unfortunately, neither public nor state-maintained records provide a dataset with the minimum measure (test and student group) used to determine whether a school failed or not, requiring the researcher to take publicly available data and apply each of the described provisions to all potential school measures to find the minimum score used to classify school failure. The researcher verified results of the application of these provisions by comparing calculations to 250 randomly sampled PDF and html documents describing individual school ratings.

[20] Calculated numbers regarding changes to the minimum score used publicly available campus-level test performance data.

[21] Required Improvement (and some other provisions) not only shielded a minimum score from consideration. Depending on the year and the standard, RI could also move a campus from acceptable to recognized or from recognized to exemplary status. The discussion in this section specifically focuses on the application of various provisions to the minimum score regardless of their rating, which may have been above passing and so may include schools that would have passed without the provisions.

changed from being reading for economically disadvantaged students to being math for economically disadvantaged students. After the application of RI, affected campuses experienced an average increase of 5.24 percentage points in their 'minimum' passing rate.

For two years, 2009 and 2010, Texas' accountability system included a student growth measure called the Texas Projection Measure (TPM). If an individual student's learning grew more than expected from the previous year to the current year then the student 'passed' that measure. Once TPM re-calculated each student's passing status, it then re-calculated the passing rate for each group using this new numerator. TPM calculations often altered the previously determined minimum score; in 2009 more than 70% of campuses applied TPM to their minimum score, and in 2010 this occurred for more than 78% of campuses across Texas. For 40% of these changes, the TPM revised figure modified the minimum score for the same test for the same specific population of students; so a school with a minimum score for Hispanic students in mathematics was improved to passing. A further 19% of minimum scores altered through TPM changed the specific population of students at the minimum, but retained the same type of test. Another 18% of TPM modifications changed the type of test that was the minimum, but did not change the specific population of students at the minimum. For the two years of the TPM, affected campuses increased their average minimum score by 7.66 percentage points.

Schools with many measures (tests and specific populations of students) received additional help because depending on the number of measures, Texas excluded some from consideration. If a campus needed to pass four or fewer measures, no exceptions could be applied. If a campus needed to pass between five and eight required measures, the campus could employ one exception. Between nine and 11 required measures provided the campus two exceptions; between 12 and 15 required measures provided three exceptions; and, for campuses with more than 16 required measures, four exceptions. Unlike other provisions,

campuses could not apply the same exception in consecutive years. The Texas Education Agency applied exceptions, just as with RI and TPM provisions, automatically. Exceptions impacted the minimum score for between 1% and 3% of schools, depending on the year; however, during the years TPM was applied, only 80 campuses applied exceptions to their minimum score. Half of campuses that applied exceptions that changed the minimum score changed the specific population of students at the minimum, but retained the same type of test. Just under a quarter (23%) of campuses where exceptions changed the minimum score, this change altered the type of test at the minimum, but saw the same population of students at the minimum. Campuses applying exceptions experienced an average increase in their minimum score of 8.19 percentage points.

Additional provisions also influenced some schools' classification. Two of these special provisions specifically focused on the minimum score: the school leaver, and the federal racial/ethnicity provision. A third provision dealt with small sample sizes. The school leaver provision in 2007 and 2008 prevented completion and dropout rates from being the sole reason a campus could fail. Thus, if all other measures for the school remained above the passing threshold, but the school's dropout rate failed to meet state requirements, the school would pass.

The collection of race and ethnicity information in Texas changed to match the federal definition in 2010. In previous years, students in Texas chose a single categorical choice for their racial and ethnic characteristics; beginning in 2010, students could elect to choose multiple races and choose their ethnicity separately. For students who took the TAKS test in 2011, the federal race/ethnicity provision linked student responses about race and ethnicity from the previous testing administration and then recalculated the school's performance based on both the old racial/ethnic record and the new racial/ethnic record. If

a student group at a school passed based on their previous racial/ethnic classification, that group met the standard.

Finally, for schools with fairly small populations, the Texas Education Agency engaged in a special analysis to determine passing or failing rates and the school's rating; this special analysis occurred in each of the study years and relied on generally unclear methods. School rating information indicates a school's passing status and that TEA applied some special analysis to determine that rating but no other information.

Even after TEA applies all these provisions, a school may still appeal the rating. The accountability handbook clearly states that schools should only do so under limited circumstances. If after applying all provisions, exceptions, and appeals, the minimum passing rate of students from any of the interested populations (assuming sufficient size) on any of the five tests is below the required passing threshold that school has failed. Once a schools fails for two years in a row, the school must reconstitute. Neither the test nor the student group at the minimum passing rate need to be the same, so a school where Hispanic students fail to meet the passing threshold in mathematics in the first year and White students fail to meet the passing threshold in reading in the second year must reconstitute.

The Threat and Implementation of Reconstitution

Schools that fail immediately face the threat of sanction; if the school fails the following year, it must reconstitute. Schools under the threat of reconstitution need immediate year-to-year gains in their minimum passing rate to avoid reconstitution. Schools required to reconstitute have a year to plan prior to reconstituting in the summer after the planning year. Performance during the planning year does not influence whether the school will reconstitute or not.

While large diverse schools with the option of many exceptions likely benefit from certain shielding provisions, the obtuse structure and application of these provisions applied by the state hint at the difficulty of gaming the system. Furthermore, as the passing threshold changes from year to year, student groups passing at the threshold of any subject might fail in subsequent years with more stringent passing rate requirements. Another difficulty in managing student failing using state standards lies in correctly identifying which students require assistance. As federal accountability provisions require the same student group to fail in consecutive years in order to place the school into the next 'stage' of accountability sanctions, school staff may choose to focus on the 'failing' student group from last year to the detriment of other student groups. Texas' accountability standards prevent this type of focus since failure by any group on any test may lead in any year to school failure to meet the standards. Thus the trigger system, while being clear, presents enough complexity that poorly performing schools entering into the testing season likely cannot accurately predict which group might fail, and, even if able to do so, are unlikely to know whether that failure will, after the application of various provisions, actually cause that school to be classified as academically unacceptable after the application of minimum passing score provisions.

Data Structure and Variables

This panel data includes information from 2003 through 2011 on students attending and matriculating through campuses, with both student-level information and aggregate campus-level information. The final dataset brings together information within the Texas Education Research Center and publicly-available datasets through the Academic Excellence Indicator System. Staff records, including the types of staff present on a campus such as teachers, principals, and counselors, provide detailed information for the formation of campus-level aggregates of human capital measures and faculty and student

racial balance measures. Course information provides evidence of changes in academically rigorous coursework taught such as Advance Placement, International Baccalaureate, and magnet courses. Publicly available records provide information on the academic classification of a school determining whether the school potentially faced or executed reconstitution. This work employs two datasets for analysis, one set examining school-level outcomes and another for student-level analysis.

Variables for School-Level Results

Dependent Variable

The second research question (Does reconstitution help failing schools improve their minimum passing rate?) utilizes the campus-level minimum passing score. This minimum passing score takes the state standard required passing rate as seen in Figure 4.1 and subtracts it from the minimum passing rate for that campus as described above. Each test in each year possesses a unique mandated threshold. Shifting the minimum passing rate to the minimum passing score effectively places schools that fail to meet state standards below zero and schools that successfully meet these standards above zero.

Attrition

A wider issue in any research design of involves attrition (absence of the outcome or dependent variable) for some units within the data. Attrition may rob any study of its validity and generalizability (Shadish 2002). Attrition occurs in both school-level and student-level data. Outcomes for schools may be absent from the data for two reasons: school closing prior to outcome measurement outcomes (this occurs four times) and truncation of the dataset after the year 2011 (this occurs 27 times).[22] Many instances of

[22] Assignment to the cause of attrition is as follows: schools that trigger reconstitution in 2010 or later are classified as 'truncated'; non-truncated schools with no record at all in the dataset of any outcome the year the school was slated for reconstitution are classified as 'closed.'

closing (three of four) or truncation (thirteen of 27) occurred at previously reconstituted schools. Truncation of the data is a necessary feature of this work as state-mandated tests change prior to and immediately after the period examined; tests on either side of this period are not comparable to the tests taken in between.

Independent Variable

The independent variable determines whether a threatened school will be reconstituted and so acts as a forcing variable. This forcing variable is the minimum passing score (described above) for the school the year after it failed for the first time; schools with a minimum passing score above zero will not reconstitute and schools with a minimum passing score below zero will reconstitute. Specifically because it assigns treatment this independent variable is used to examine the second research question. Due to the nature of the chosen method of analysis for the second research questions (regression discontinuity) the analysis lacks any additional independent variables.

Variables for Student-Level Results

Dependent Variables

The third research question and its corollaries employ modified individual student performance on math and reading TAKS tests as the dependent variable. Since tests changed across grades, years, and subjects, student raw test scores were standardized by subtracting their yearly subject grade mean and divided by the standard deviation of the raw score. Hence, the dependent variable for all students in the 8^{th} grade in 2008 for math included each student's raw math test score subtracted by the mean of all raw math test scores in 8^{th} grade math in 2008, divided by the standard deviation of the raw test scores in 2008 8^{th} grade. This linear standardization of student test scores results in a z-score and

allows for a reasonable comparison of student performance, relative to all other students in the state, between years and for individuals across years.

Independent Variables

The third set of research questions apply a broad set of independent variables, including the primary variable of interest (whether and when a student attended a school immediately after its reconstitution), a set of potential mediator variables (which theory suggests may explain any effects of reconstitution), a group of moderator variables (which our hypotheses suggest might strengthen or weaken the relationship between reconstitution and student performance), and, finally, a set of control variables (intended to account for some student and campus variation).

Variable of Interest: A binary variable indicates whether a student attended a school the year it reconstituted (one) or not (zero). Longitudinally, for students who attend a school the year it reconstituted, this binary variable is one during and after attending the school and zero prior to attending the school. For students who never attend a reconstituted school, the binary variable is zero.

Mediating, Moderating, and Control Variables: The theoretical framework provides the basis for a number of measures which may be behind the influence of school reconstitution. Many of these variables come directly from available administrative data, requiring little modification. For variables created from administrative data, a section discusses the process of their construction.

Mediating Variables: Mediating variables relate primarily to campus human capital and racial balance hypotheses. Human capital variables include multiple measures of staff and teacher capital on the campus. This includes the share of staff (teachers, principals, assistant principals, and counselors) to that campus that year, as campuses are

required to replace staff when they reconstitute. The share of principals and teachers with graduate degrees and an indicator of their experience provide standard measures of their human capital. An additional human capital variable is the mean number of years that teachers worked with the current principal. Administrative datasets contain these measures and these are aggregated to the school-level for each year.

Measuring racial balance requires more finesse. Traditional measures of racial isolation aggregate smaller units into larger units to compare the distribution of racial characteristics (Sakoda 1981). These methods describe isolation as racially isolated classrooms within a school or isolation at a school compared to the rest of the school district. While helpful, these calculations do not provide direct comparisons between two groups, but a group within a larger group. Euclidean distance (Formula 1) provides a relatively easy means through which to measure distance between two vectors, X and Y. In this case, each vector includes the shares of Hispanic, African American, White, and Asian students; each of these racial/ethnic groups is represented by a subscript of 1, 2, 3, or 4 respectively. The vector Y contains campus-level shares of these students for each year and the vector X contains the same shares for a comparison set of percentages, for example, the averages across Texas.

$$D=\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2}$$

As most schools have shares of students from any particular demographic of less than 100%, such a calculation will ensure that nearly all distance measures, regardless of how they are specified, will be between zero and one. Thus, a distance measure of zero will indicate that a school is perfectly racially 'balanced' between teachers and students, and a distance measure of one (or greater than one) indicates the school is notably lacking in racial balance between teachers and their students.

Two other sets of mediating variables relate to how failing campuses might allocate their resources to improve their performance outside of the standard process of reconstitution. Schools might focus on specific groups of students needing help. If campuses focus on students from the group with the minimum passing score last year, the school's focus on this group might explain any effect of reconstitution on student achievement (RQ3.2). Another mediating variable is a dichotomous variable indicating whether a student failed the subject-specific test (RQ3.3) that was responsible for the minimum passing rate from the previous year. Finally, a collection of variables describing advanced or innovative courses, which campuses might choose to implement as part of the reconstitution process. These mediating variables describe those features intended to explain any effects of school reconstitution on student achievement.

Moderating variables: The effects of reconstitution might differ based on student factors, such as race and ethnicity; these variables moderate the estimate of the variable of interest and the outcome, here test performance. Earlier discussion of the research questions identified several moderating variables. A set of variables describing the racial/ethnic, gender, and FRL status of the students (RQ3.4) allows for the investigation of whether reconstitution provides greater benefit for specific groups of students.

Control Variables: A number of control variables account for time-varying general student and campus characteristics. These include the number of campus staff and whether the school has instituted new curriculum, such as Advanced Placement, International Baccalaureate, or magnet courses. Statistical models (discussed below) allow for control of many individual-level characteristics such as underlying academic ability, and willingness to put forth effort.

Description of Available Data:

Research applied two separate datasets: that for campus-level analysis and that for student-level analysis. Campus-level analysis for the regression discontinuity design solely utilizes data derived from publicly-available information published by the Texas Education Agency.[23] Student-level analysis includes both individual administrative data from the Texas Education Research Center and publicly available school-level data. The set of variables used in the analysis is described in Tables 2 and 3.

Individual-level standardized math and reading scores possess a mean of zero and a standard deviation of one. Fixed effects estimates of the influence of reconstitution on student performance thus represent the aggregate change in the standard deviation for students who experience reconstitution. Positive values hint at increases in academic performance relative to their peers and negative values describe declines in academic performance relative to their peers. Estimates represent the differences in performance over the life of the student for all years after the student experiences reconstitution; results do not differentiate by grade level.

[23] The reasons for doing so include testing whether and to what extent this policy is capable of being evaluated using information already provided.

Student Variable	Obs.	Mean	Std. Dev.	Min	Max
Math Score	25788482	0.000	1.000	-5.555	3.521
Reading Score	26127434	0.000	1.000	-6.221	3.150
Student After Reconstitution	26261892	0.005	0.072	0	1
Student at Future Reconstituted School	26261892	0.008	0.086	0	1
Student Experienced Reconstitution	26261892	0.002	0.050	0	1
White	26261892	0.366	0.482	0	1
Hispanic	26261892	0.453	0.498	0	1
African American	26261892	0.141	0.348	0	1
Asian	26261892	0.033	0.178	0	1
Ever FRL	26261892	0.607	0.488	0	1
Male	26261892	0.509	0.500	0	1
Special Education	26189835	0.102	0.302	0	1
Gifted	26211560	0.095	0.293	0	1
At Risk	26214018	0.439	0.496	0	1
Student Years on Campus	26261892	1.656	0.826	1	9
Student Years on Campus (Adjusted)	26261892	1.654	0.817	1	5
Part of Group with Lowest Passing Rate Last Year	26261892	0.348	0.476	0	1
Failed Math Test Last Year	26261892	0.144	0.351	0	1
Failed Reading Test Last Year	26261892	0.279	0.448	0	1
Year	26261892	2007.165	2.585	2003	2011

Table 4.2 Means of Student Variables

School Variable	Obs.	Mean	Std. Dev.	Min	Max
Total FTEs of Teachers	26208479	69.418	48.017	0.028	310.400
Total FTEs of Principals	25396266	1.021	0.283	0	7
Total FTEs of Counselors	24826926	2.827	2.498	0	16.884
Total FTEs of Assistant Principals	22685415	2.585	1.959	0	16
Total FTEs of Other Staff	26064759	15.028	8.877	0	153.237
Share of Principals with a Master's	25396266	0.878	0.323	0	1
Principal Tenure	25396266	3.992	2.970	1	16
Share of Teachers with a Master's	26208479	0.227	0.101	0	1
Teacher Tenure	26208479	4.937	1.588	1	16
Share of Teachers New to Campus	26208479	0.223	0.157	0	1
Share of Teachers One Year on Campus	26208479	0.159	0.115	0	1
Share of Teachers Four or More Years on Campus	26208479	0.067	0.065	0	1
Share of Principals New to Campus	25394641	0.224	0.410	0	1
Share of Principals One Year on Campus	25394641	0.192	0.390	0	1
Share of Principals Four or More Years on Campus	25394641	0.081	0.270	0	1
Vector Distance	25394741	0.539	0.209	0.006	1.719
Mean Number of Years Experience with Principal	25318952	2.040	1.263	1	16
Any Innovative Courses	25396262	0.019	0.136	0	1
Any Magnet Courses	25396262	0.030	0.169	0	1
Any AP Courses	25396262	0.332	0.471	0	1
Any IB Courses	25396262	0.019	0.136	0	1

Table 4.3 Means of Campus Variables

Additional samples discussed below also present means of variables.

STATISTICAL METHODS

Several methods attempt to determine the effect of reconstitution on student achievement. First a demographic analysis of schools across Texas and reconstituted schools looks for patterns of change related to school reconstitution. Next, a regression discontinuity design utilizes school minimum passing scores to determine whether reconstitution successfully helps failing schools improve their minimum passing score. Finally, a fixed effects model using student-level data examines whether students who

attended a school immediately after it reconstituted experienced sustained improvement in their math and reading achievement. This final analysis will include an examination of both moderating and mediating variables, with the goal of establishing which theories explain any effects of reconstitution (Baron 1986).

Demographic Analysis: The relatively large number schools reconstituted in Texas permits aggregate information related to the shares of staff new to the school in any given year, whether teachers at reconstituted schools possessed more education or experience than in previous years, and if the faculty make-up more closely matches that of their student body than in previous years.

Regression Discontinuity: Schools that fail to meet state standards for the first time are in immediate peril of being reconstituted. If their minimum passing score does not reach the passing threshold they must reconstitute. The minimum passing score the year after schools fail the first time solely determines whether a school reconstitutes or not. One of the chief difficulties in evaluating the effect of a policy on student outcomes where students are assigned to a treatment or control group based on test scores stems from the many tests a student takes, which can greatly complicate efforts to use the regression discontinuity design (Reardon 2010). Here, all students within a school are assigned into a treatment or control group based on the minimum passing score of a student group at that school, eliminating this complication.

Schools whose minimum passing score falls below the standard experience reconstitution while schools whose minimum passing score lies above the standard do not. Moreover, the minimum passing score not only serves to determine which schools receive the treatment, but also serves, in later years, as the outcome variable of interest. The assignment being completely determined by the minimum passing score, and the expectation that the conditional distribution would be continuous and uninterrupted across

the value of the predictor (the minimum passing score) absent the intervention indicates that a regression discontinuity design may be used to establish the causal effect of reconstitution (Imbens and Lemieux 2008).

This regression discontinuity design is complicated by the structure of the intervention itself (Figure 4.2). Schools become eligible for reconstitution only after failing to meet the minimum passing threshold (Year 1). Hence, non-failing schools are excluded from (or filtered out) of the analysis.[24] Once eligible for reconstitution, the school's minimum passing score in Year 2 determines whether the school will be reconstituted, acting as a forcing variable: schools that do not meet the standard that year will eventually reconstitute, whereas passing schools will not. Schools classified as failing in Year 2 take all of Year 3 to plan for their reconstitution immediately prior to Year 4.[25] The minimum passing score in Year 4 serves as the outcome variable for reconstituted schools. For schools classified as passing in Year 2, and so do not reconstitute, their minimum passing score in Year 3 serves as their outcome variable.

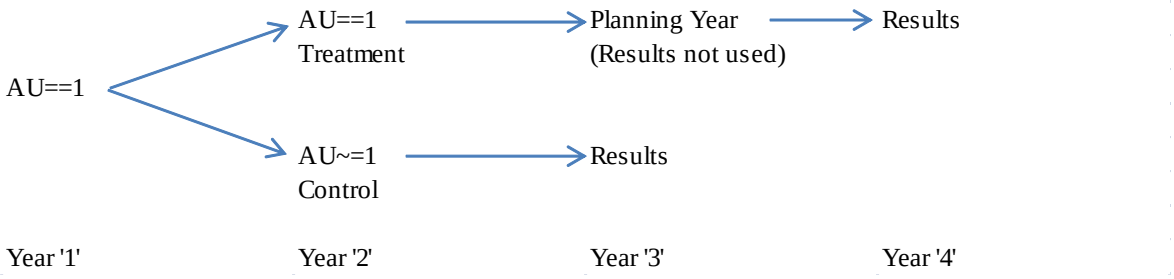

Figure 4.2 Structure of Relevant Performance for Regression Discontinuity Designs

[24] As reconstitution purposely targets failing schools, removing non-failing schools from this design is appropriate.

[25] Testing results in Year 3 can neither 'save' the school from reconstituting that summer through excellent performance nor worsen the sanction imposed. This distinction will play an important role using student-level data later on, as students who arrive at a school during the planning year do not play a role in designating the school as failing but do experience reconstitution.

Using the Figure 4.2 as a guide, the description of the regression discontinuity design is as follows. For every school (*i*) in each year (*t*) $Y_{i,t}$ indicates the outcome variable (in this case the minimum passing score at the school for year t), and $X_{i,t}$ represents the forcing variable which determines whether or not a school is assigned to be reconstituted. Henceforth, for the sake of convenient notation, year t represents the year when schools are assigned or not to reconstitution (Year 2 above). Assume that $Y_{i,t+2}(1)$ and $Y_{i,t+1}(0)$ indicate the minimum passing score for reconstituted schools and schools under threat but not reconstituted. Any single school is required to reconstitute if their minimum test score threshold in Year 2 ($X_{i,t}$) is less than zero and all schools above the minimum test score threshold in Year 2 do not reconstitute. The observed outcome in this case is the following.[26]

$$Y_i = \begin{cases} Y_{i,t+2}(1) & \text{if } X_{i,t} < 0 \\ Y_{i,t+1}(0) & \text{if } X_{i,t} \geq 0 \end{cases}$$

This cutoff provides the opportunity to estimate the average treatment effect (τ) at the cutoff value where τ represents the difference between the limit of the estimated fitted function as it approaches the cutoff value from the left (μ_+) and the limit of the estimated function as it approaches the cutoff value from the right (μ_-). Thus, the difference in the expected values of the estimated fitted functions on either side of the cutoff at the cutoff value (the limits of each respective function) represent the average treatment effect.

However, minimum passing scores might not be expected to jump at the cutoff value if reconstitution improves school performance more for schools with lower minimum

[26] As statistical models in this section lead to results using Stata's rdrobust, model specification use notations from articles from the authors of this statistical routine (Calonico 2014). Note that the form of this function, by specifying time t, allows the assignment of multiple outcomes for schools placed under threat multiple times.

passing scores and less for schools near the passing threshold. This scenario describes a noticeable shift in the slopes of the estimated functions on either side of the cutoff value, referred to as a kink regression discontinuity (introduced by Nielsen, Sorensen, and Taber (2009)). Kink regression discontinuity measures the difference in the derivative of the underlying estimated functions at the cutoff value (Calonico, Catteneo et al. 2014). Whereas graphical evidence of standard regression discontinuity would show a break between the two estimated regression functions at the cutoff value a kink regression discontinuity would show an angle.

All regression discontinuity designs make specific assumptions when calculating the confidence intervals of any treatment effect estimates. Recent work by Calonico (2014) provides robust confidence intervals based on bias-corrected local polynomials. These robust confidence intervals are more accurate than conventional calculations when using relatively large bandwidth choices, offering calculation of robust bias-corrected confidence intervals for average treatment effects (Calonico 2014).[27]

While all types of regression discontinuity designs place fewer required assumptions on their form and interpretation than other non-experimental techniques, graphical evidence does not alone ensure statistical significance (Lee and Lemieux 2010). Scholars note that regression discontinuity designs may be invalid if their subjects are able to precisely manipulate the outcome variable and have developed methods to determine empirically whether this occurred (McCrary 2008). Such non-random sorting to receive or (in this case) avoid treatment is a logical consequence of this policy, as demonstrated in the density of the forcing variable ($X_{i,t}$) in Figure 4.3 (McCrary 2008).

[27] Conventional, bias-corrected, and robust estimates produced in this dissertation rely on CCT's Stata routine rdrobust.

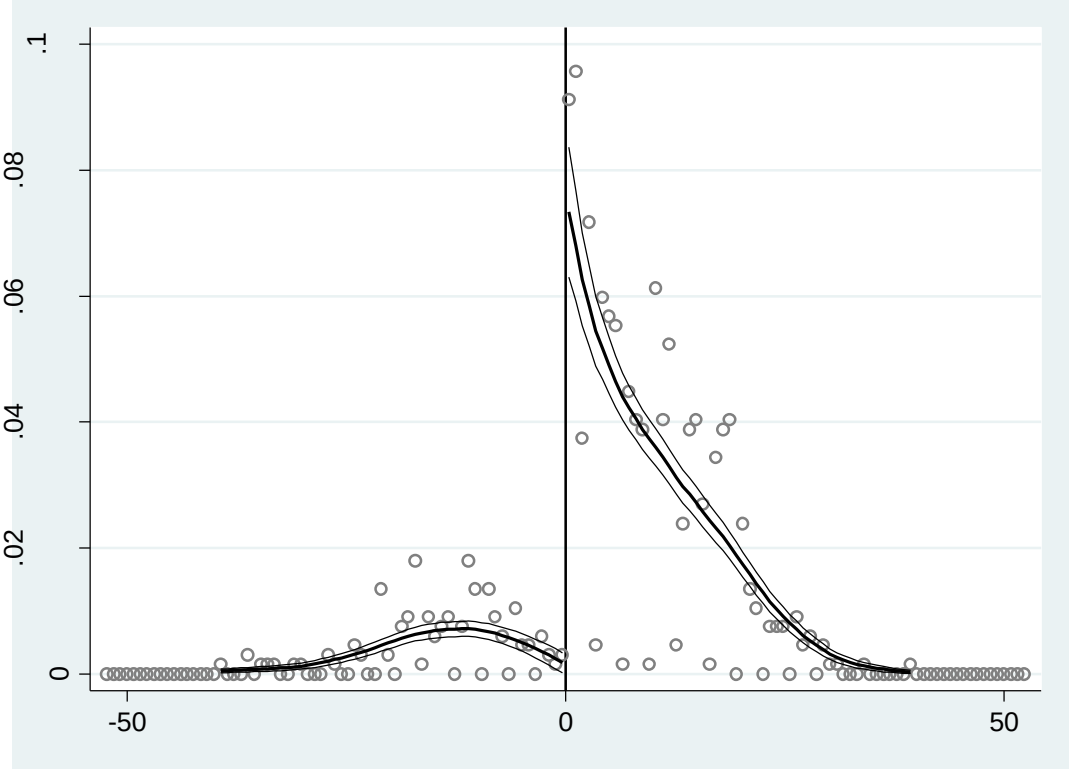

Figure 4.3 Density of Reconstitution Assignment Variable

The leap in the density of schools at the cutoff value provides evidence of the desire to avoid reconstitution. After failing for one year to meet state standards, staff at schools possess incredible incentives (e.g., job loss avoidance) to ensure the school does not fail the next year. Schools with the capacity to improve student learning and the ability to both successfully monitor student needs and provide related services will engage in the effort needed to move their minimum passing score above zero the following year. This also hints that schools below the passing threshold lack such capacity for rapid improvement. As the manipulation of the outcome variable stems from avoidance and schools who fail to avoid reconstitution inherently lack the capacity for rapid improvement, treatment schools are those in most need and face the greatest barriers. That being the case, any estimates of the effect of reconstitution are likely biased downward, since schools that

narrowly avoided reconstitution possessed the inherent capacity for rapid improvements while their failing compatriots lacked this capacity.

Sensitivity to model specifications for estimating the average treatment effect include the choice of bandwidth (the width of the window of observations used for the local linear regression) and the kernel (which provides weights for the local linear regression in relation to the cutoff value) (Lee and Lemieux 2010). While optimal methods to determine the bandwidth are available (Imbens 2009), these bandwidth options and other bandwidth choices are considered and discussed in the results section. Initially estimated and presented results utilize a uniform kernel (no weighting of units related to their distance from the cutoff value) and alternative estimations using other kernel choices are presented and discussed.

Presented initial estimates include all reconstituted schools in all years; thus, data includes schools reconstituted more than once with their related results. Additional sensitivity analysis considers whether including or excluding schools reconstituted more than once changes the results of the analysis.

Fixed Effects Models: While the policy clearly intends to improve school-level performance, school-level outcomes differ from the larger question of whether reconstitution engenders improvements in student performance. Students attending and testing in a particular year at a school are not the same students as in the previous year. In each year, a portion of students matriculate out of the school (e.g., eighth graders leave middle school and go to high school as ninth graders), and another portion of students matriculate into the school (e.g., rising fifth graders enter middle school as sixth graders). Additionally, students may newly enroll in a school if their parents move into a different catchment area. Changes in school-level aggregate student outcomes may reflect student mobility and matriculation. Second, students do not take the same test from year to year,

even if they reside at the same school (e.g., students who completed the third grade test last year will complete the 4th grade test this year). While some cohorts of students may have difficulty at every test they take while at their school, other cohorts might not, which could influence the passing rate, particularly for tests that are only intermittently taken, such as science and social studies. Thus campus-level analysis neglects to answer the larger and, perhaps more important policy question: does school reconstitution improve student-level achievement?

If students are randomly assigned to treatment (reconstitution) and control schools then estimates of treatment (being uncorrelated with the error term) would provide an unbiased estimated using just Ordinary Least Squares (OLS). Students are not randomly assigned to treatment as schools draw their population from assigned catchment areas, thus treatment is likely correlated with the error term. The conditional independence assumption implies that if all observable characteristics are included in the OLS model, then the treatment is uncorrelated with the error term and estimates of treatment effect are unbiased. Including some but not all observable characteristics into a model leaves it open to critiques of omitted variable bias (OVB). If the omitted variable is correlated with both the treatment and the outcome then estimates using standard OLS will not provide true estimates of treatment effect. As the panel data provides multiple observations about students over time, time-invariant unobserved and observed characteristics may be removed by subtracting student-level averages from both sides of the OLS equation; doing so is equivalent to estimating a fixed-effect model. A key advantage of using a fixed-effect model in evaluating education policy involves the many unobservable, but likely fixed student characteristics including capacity for working toward learning, underlying academic ability, family make-up and background, individual learning preferences, and year-to-year stable peer-networks, are all effectively accounted for in the fixed effect

model. In effect, each student then serves as his or her own control group in any fixed-effect analysis (Angrist 2009).

The two-way fixed-effects model applied in this dissertation is of the following form:

$$Y_{it} = \alpha_i + \rho D_{it} + \beta_1 X_{it} + \beta_2 Z_{ct} + \gamma_p + \delta_t + \varepsilon_{it}$$

Where Y_{it} is the dependent variable (DV) for student (i) at time (t); D_{it} is a binary variable representing whether student (i) received treatment at time (t), which describes the multiplicative interaction between treatment and the post-treatment period, so ρ is the estimate of the effect of treatment. The variable α_i is the unknown intercept for each student. X_{it} represents (in this case) a vector of independent variables where β_1 describes the coefficient for each IV. The variable Z represents a vector of campus-level variables for campus (c) at time (t), with β_2 being its coefficient; Note that students are rarely confined in the records to a single school as they transition across grade levels and transfer between schools. The γ and δ terms represent the coefficients for dummy explanatory variables (i.e., the student fixed effects on test performance and the year fixed effects on test performance, respectively), while ε_{it} is the error term.

Prior to implementing any fixed effects models, several preliminary statistical tests are applied. First, to ensure the appropriateness of a fixed effects model as opposed to a random effects model, a comparison test of these models (a Hausman test) determined that the error terms are not correlated with the regressors. Then, an F-test determined that dummy variables for each year were not jointly equal to zero; thus time fixed effects are

appropriately included in regressions.[28] All fixed effects models employ robust standard errors to account for any idiosyncratic patterns or heteroskedasticity in the data.[29]

The fixed effects model specification is applied numerous times using different individual student and aggregate campus variables as well as different samples of students. This section describes the samples of students used in the analysis and is followed by a brief discussion of the process of examining moderating and mediating variables in relation to school reconstitution. Several extensions of this analysis involve making slight changes to which students are classified as having experienced a reconstituted school.

Three samples of students are used in fixed effects estimates. First, results are calculated including all students in Texas. Second, only students who ever attended a threatened or reconstituted school are included in the analysis. Students who attended a reconstituted school the first time during the planning year and students who attended the reconstituted school the fall after the summer it was reconstituted compose the third sample. As necessary, moderating and mediating results are presented for each of these samples. Discussion here parallels the structure of the results chapter, with sections specifically related to the sample used in the analysis.

All Students in Texas and Students Attending Failing Schools in Texas

Prior to estimating whether reconstitution improves student achievement an appropriate context needs to be establish which should provide clearer meaning of any future estimates. The structure of the variable of interest in the fixed effect analysis implies

[28] Alternate attempts at incorporating campus or district fixed effects estimates were attempted, but the computational needs for their implementation were too great. Attempts to use even a limited sample of students (those who ever attended a threatened or reconstituted school) to test whether these fixed effects were necessary proved too large.

[29] Ideally, standard errors would employ clustering at the campus-level. However, students do not remain fixed within schools and so calculation of campus-clustered standard errors is not possible using fixed effects models within the statistical software package used (Stata) for this research.

that students who attend a reconstituted school the summer following its reconstitution experience a long term effect over the course of their academic career in public schools.

To gauge the meaning of calculated estimates, a series of 250 simulations (performance twice, once for math and once for reading) using state data provide a set of estimates of the general 'effect' of attending a set of randomly selected schools. These simulations begin with the random selection of 15 schools for each year. These selections occurred without replacement and could include any school in Texas. After all the schools were selected, students attending these schools in that year were assigned as receiving 'treatment' and students who did not were assigned as 'controls.' Then estimates of the effect of receiving 'treatment' were calculated for both math and reading tests. Performing these simulations allows readers the opportunity to examine a distribution of estimates for schools in Texas that may or may not have been reconstituted. The distribution of the estimates of these simulations provide some indication of where the effect of attending a reconstituted school (both prior to and after reconstitution) falls on the simulation distribution. Since little to no estimates of the effect of reconstitution on student achievement exist in the literature, these efforts are intended to provide context to estimated results.

After establishing general context, initial estimates of the effect of reconstitution include all students in the state of Texas. These results, along with estimates from students who attended the school two years ago are placed within the simulated estimates distribution. Table 4 provides the means of students who ever attended a reconstituted school in relation to means of students who did not, demonstrating that these two groups are statistically dissimilar.

	Treatment		Control	
	Mean	Std. Dev.	Mean	Std. Dev.
White	8.37%	0.277	37.03%	0.483
Black	32.80%	0.469	13.86%	0.346
Hispanic	57.59%	0.494	45.09%	0.498
Asian	0.76%	0.087	3.33%	0.179
Male	51.25%	0.500	50.94%	0.500
Ever FRL	91.36%	0.281	60.25%	0.489

Table 4.4 Means of Select Variables for Treatment and Control Groups in Overall Analytical Sample[30]

However, reconstitution is primarily intended to improve the performance of failing schools. The second analysis sample includes only students who either attended a failing school or attended a reconstituted school the year after it reconstituted. These estimated effects align with the analysis and sample presented in the regression discontinuity design in the previous section. Note that the characteristics of students who attended threatened schools still differs from the characteristics of students who attended reconstituted schools.

	Treatment		Control	
	Mean	Std. Dev.	Mean	Std. Dev.
White	8.37%	0.277	22.40%	0.417
Black	32.80%	0.469	22.94%	0.420
Hispanic	57.59%	0.494	52.97%	0.499
Asian	0.76%	0.087	1.16%	0.107
Male	51.25%	0.500	50.66%	0.500
Ever FRL	91.36%	0.281	78.56%	0.410

Table 4.5 Means of Select Variables for Treatment and Control Groups in Threatened or Reconstituted Sample

[30] Not all model variables are included for the purpose of brevity to demonstrate major differences between treated and non-treated individuals in the sample.

To determine if students from certain backgrounds might see more or less benefit from reconstitution, the sample is limited to include only these populations; since these characteristics are time invariant, limiting the sample is the only means of providing an estimate of any differential effects. Potentially mediating covariates are included in various stages to determine whether they attenuate the effect of school reconstitution on student achievement.

One effort to test the sensitivity of results uses a modified definition of who received treatment when. This sensitivity test looks at whether removing students matriculating to a reconstituted school the year after the summer it reconstituted changes the estimates of the effect of reconstitution. Students who arrive at a school immediately after it has been reconstituted do not actually 'experience' reconstitution and so are placed, in this part of the analysis, in the control group.

Special Analysis

One problem with using student-level fixed effects models to demonstrate a causal relationship between reconstitution and student performance stems from issues related to endogeneity. Students who attend a reconstituted school might receive benefit from its reconstitution, but their aggregate performance led to the school's failing status and, eventually, reconstitution itself. This is one of the reasons using alternative methods which traditionally lead to more accurate causal treatment of an intervention, such as propensity score matching or weighting, is not applied here. Any matches or weights linking individual students who themselves 'cause' reconstitution are problematic. Fortunately, how reconstitution is applied allows for a more rigorous analysis.

Recall that reconstituted schools fail two years in row, engage in a planning year for reconstitution, and then reconstitute over the summer prior to the next year. School-level student outcomes during the planning year cannot prevent reconstitution and so are

not tied to treatment assignment. Hence, the test performance of students whose first year at the reconstituted school is the planning year is unlinked to whether the school is reconstituted. Furthermore, students who arrive at the reconstituted school in the post-reconstitution year have not actually 'experienced' reconstitution: the school already underwent this process over the summer. Since these students are drawn from the same catchment areas and only a year separates them, differences in their long-term outcomes from attending the school are reasonably due to reconstitution. A comparison of their means demonstrates these two groups possess similar characteristics (Table 4.6).

	Treatment		Control	
	Mean	Std. Dev.	Mean	Std. Dev.
White	8.26%	0.275	8.74%	0.282
Black	34.69%	0.476	32.24%	0.467
Hispanic	55.77%	0.497	57.72%	0.494
Asian	0.77%	0.087	0.85%	0.092
Male	51.86%	0.500	50.69%	0.500
Ever FRL	92.15%	0.269	91.03%	0.286
Part of a Failing Group	30.64%	0.461	27.81%	0.448
Failed Math Test Previous Year	21.31%	0.410	27.40%	0.446
Failed Reading Test Previous Year	15.97%	0.366	11.58%	0.320
Special Education	13.34%	0.340	9.90%	0.299
Gifted	6.04%	0.238	8.10%	0.273
At Risk	62.93%	0.483	61.75%	0.486

Table 4.6 Means of Treatment and Control for Special Analysis Sample[31]

A fixed effect model including only these two groups of students provides a clear estimate of the effect of reconstitution on student achievement. The reduction in the

[31] The rule of thumb is that linear regression methods are sensitive to normalized differences in means between treatment and control groups greater than 0.25. The largest normalized difference in means for these variables never exceeds 0.10; thus, treatment and control groups may be considered comparable for analysis purposes (Imbens and Wooldridge 2009).

analysis sample size also allows the use of district-level fixed effects in these models, which account for time invariant district effects and serve, in part, as a proxy for neighborhood characteristics.

Limitations: All presented results possess some limitations in their function and interpretation. Limitations are described in the following order: those due to data restrictions including collection and availability, those due to computational limitations of the system used, and those dealing with the nature of reconstitution's implementation in Texas. Unfortunately, data linking students to teachers was not available for the years examined in this study. Linking students to their teachers would provide more detailed information about how students were sorted into various classrooms, allowing for more detailed information regarding whether teachers new to the school received similar students compared to teachers who worked at the school during prior years. Linking student and teacher information would also permit the construction of teacher 'effectiveness' variables, an alternative and potentially important method of measuring teacher human capital. General limitations also lie in the nature of administratively collected datasets, from which this research draws its data. Somewhat detailed information is collected on students and staff, but any qualitatively-derived information, such as changes in the structure of the school into smaller learning communities, the development of new discipline procedures, or career-based academies is simply not collected.

Using over 26 million person-year records presents computational challenges. Specifically, the large sample size and use of Stata to perform the analysis made including campus or principal fixed effects in even limited samples of students problematic. Furthermore, the structure of the data could be viewed as cross-classified, with students nested within schools only during certain years and other schools during other years;

computational requirements using Stata for this type of supplemental analysis are known for their lack of timely processing.

Finally, the nature of reconstitution in Texas means that local individuals within the school district make decisions about which teachers will be retained; their reasons are their own and cannot be adequately accounted for in this research. Despite being included in the state accountability system, science and social studies tests are not utilized in the fixed effects analysis, primarily because their intermittent nature means that progress is less easily attributable to any given year.

Chapter 5: Results

The first section of this chapter examines to what extent reconstitution was implemented in schools required to undergo reconstitution. The second section of this chapter uses a regression discontinuity design to examine whether reconstitution improved the minimum passing scores of failing campuses. The third section uses fixed effects models to determine whether reconstitution improves student academic performance.

The primary mechanism of reconstitution relies on replacing large shares of teachers. While Texas provides information on the share of teachers and their lifelong teaching experience at campuses, records indicting the share of teachers new to that particular campus are absent. Between 2002 and 2012, the share of teachers new to campus in any given year ranged from a low of 23% to a high of 28% (Figure 5.1).[32] Nearly three fourths of teachers on campuses across Texas have five or fewer years of experience on that particular campus. Around 40% of teachers on campuses across Texas have less than two years of experience on that particular campus.

[32] These are averages of averages to more closely reflect the shares of teachers with these characteristics on their campus.

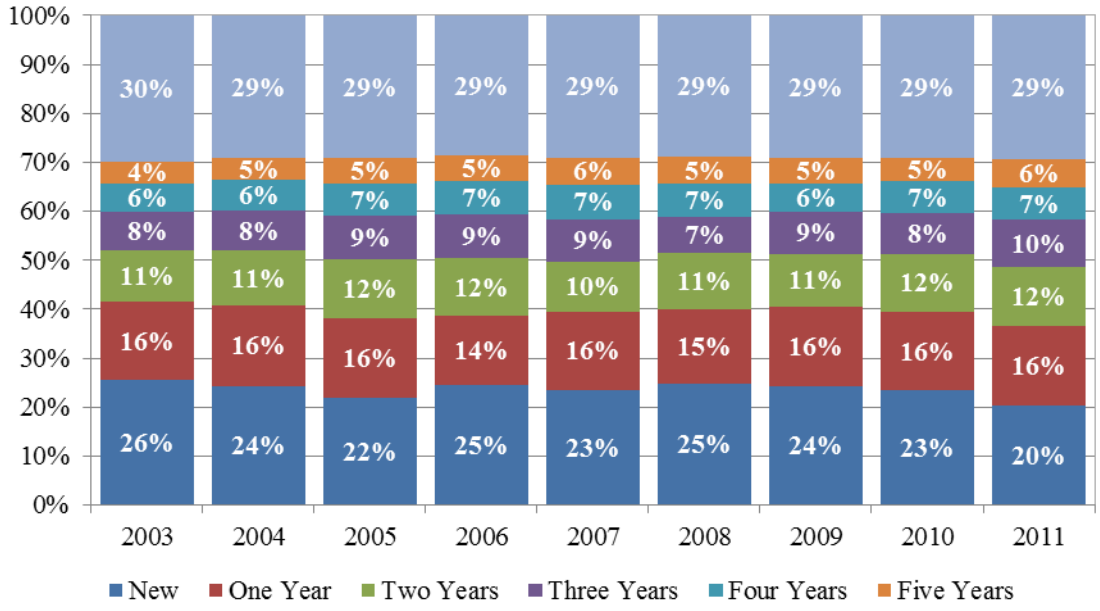

Figure 5.1 Teacher Years of Experience on Campus, by Share of All Teachers on Campus[33]

The share of teachers new to campus for schools ever reconstituted schools is consistently higher than the state averages, with rates of teachers brand new to their schools composing roughly a third of all teachers in most years and at least a quarter of all staff. Generally, these schools also exhibit a smaller share of teachers with more than five years of experience than Texas schools on average. In all but three of the years examined, more than 50% of teaching staff had been at the school for more than one year.

[33] Data presented on shares of new staff in this and other figures uses data from the Texas Education Research Center.

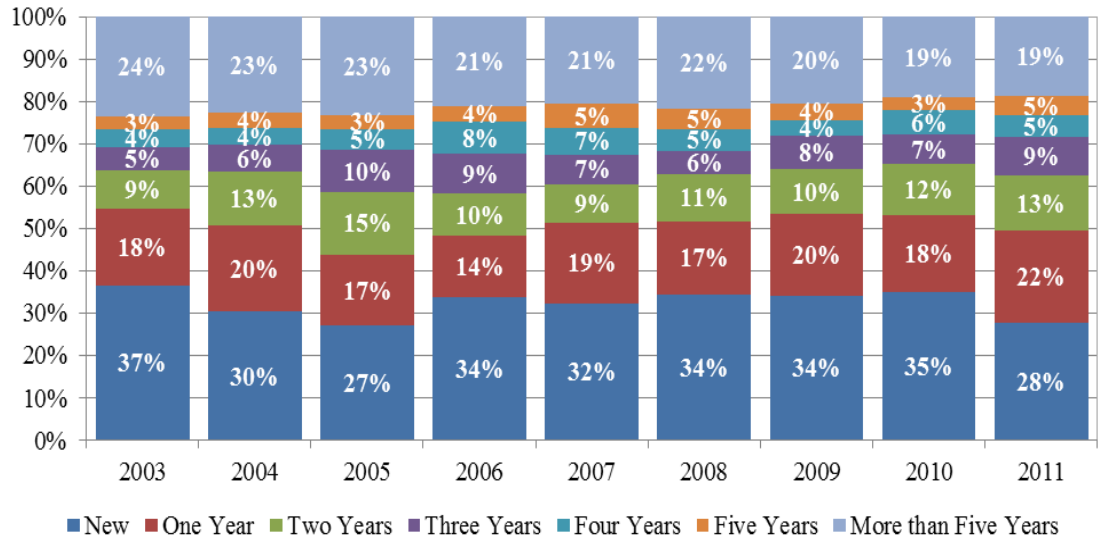

Figure 5.2 The Share of Teachers on Campus for Reconstituted Schools, by the Number of Years on Campus

A total of 137 schools were reconstituted between 2003 and 2011, during the era when the TAKS test was used for all students. The chief feature of reconstitution traditionally lies in the replacement of large shares of staff. In Texas, where individuals must re-apply for their positions and principals have broad authority in determining who stays and who leaves, the extent of staff changes are undetermined. Schools officially hear word of their academic performance rating in October, and for schools slated for reconstitution, the planning begins for reconstitution during the following summer. The reconstitution application cycle begins with a school failing any of the required measures two years in a row. The school then has a year of planning, in which their performance is monitored, though regardless of the school's performance, the school is reconstituted the next year. The share of teachers' experience on campus using this timeline reveals an increase in the share of teachers new to campus the year following being classified as being academically unacceptable for two years in a row (prior to the start of the 'planning' year).

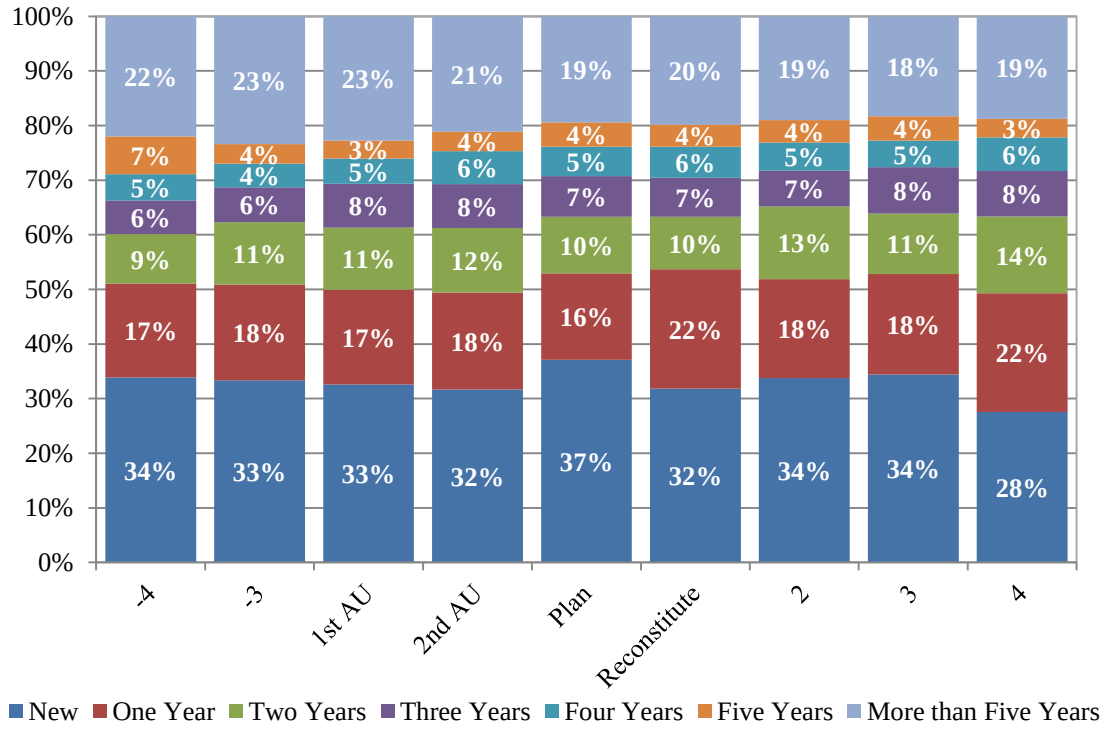

Figure 5.3 Share of Teachers with Campus Experience at Reconstituted Schools

Many teachers either choose to leave or are asked to do so during the planning year as opposed to the reconstitution year. However, teachers hired during the planning year are more likely to stay after reconstitution, with nearly 60% doing so; in other years roughly 52% of first year campus teachers remained at these schools for the following year. The pattern of the shares of teacher campus-experience leading up to reconstitution differ when compared to other types of school staff.

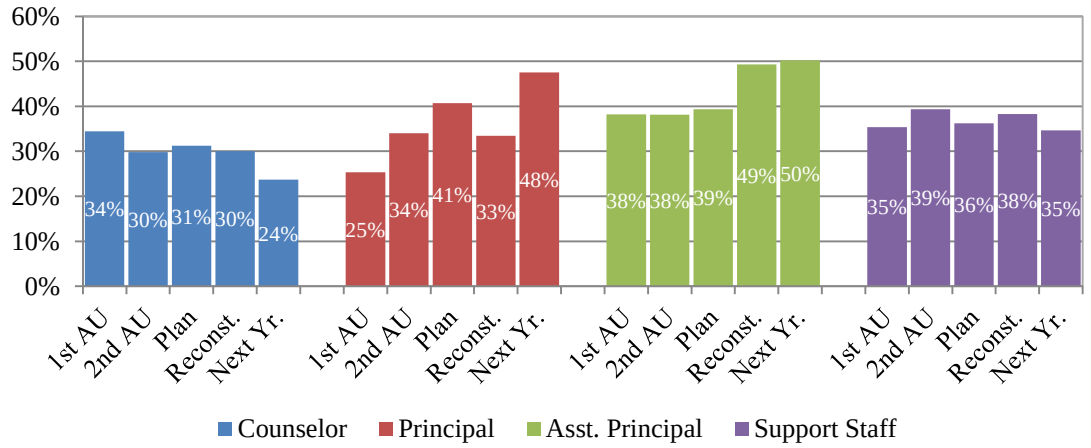

Figure 5.4 Share of Staff New to Campus at Reconstituted Schools

While the share of new counselors and support staff are relatively consistent through the reconstitution cycle, the share of principals who are new to their campus increases for reconstituted schools from the first time the campus was classified as failing through the planning year and then spikes again the year following reconstitution.

The pattern for assistant principals demonstrates that this group is the most likely to have been 'reconstituted' compared to previous years, with roughly 50% of campus-level staff being brand new to the campus during the reconstitution year. The continuing high turnover rate for the assistant principals the year after reconstitution are combined with an equally high share of principals new to their campus the year following reconstitution. Thus reconstitution in Texas unfolds without unusually high turnover for teachers, but 41% of principals are replaced prior to reconstitution.

Regression Discontinuity Results

Results of regression discontinuity analyses include the presentation of graphical evidence, regression discontinuity estimates, and a discussion of the sensitivity of results related to model and sample specifications. Graphical evidence hints at both a positive

jump in the minimum passing score for reconstituted schools in relation to their counterparts and a reduction in the slope of their estimated regression line, meaning that reconstituted schools might see improved outcomes overall and that schools with lower performance in the previous year see additional improved outcomes. Presented average treatment effects using both RD and RD kink designs include conventional, bias-corrected, and robust estimates. Changes in estimates and their statistical significance based on model specifications regarding bandwidth and kernel function selection are discussed. Changes in estimates and their statistical significance based on changes in the specification of the sample are also discussed. Conclusions regarding results are then discussed.

Graphical Results

School minimum passing scores after failing for one year assign schools to reconstitution (treatment) or not (control) based on a cutoff value of zero. Previously discussed evidence (Figure 4.3) hints at imprecise but real avoidance manipulation of the assignment variable. This sorting suggests that schools slated for reconstitution lack the capacity of other non-reconstituted failing schools to rapidly improve their minimum passing score. Graphical evidence (Figure 5.5) suggests that reconstituted schools do experience a bump in their minimum passing score in relation to previously failing but not reconstituted schools, with reconstituted schools seeing higher than expected improvements in their minimum passing scores. Figure 5.5 also suggests that reconstituted schools experience a shift (or kink) in the slope of their linear estimate in relation to non-reconstituted schools, with reconstituted schools with lower passing rates seeing more benefit than would be expected in relation to the comparison group.

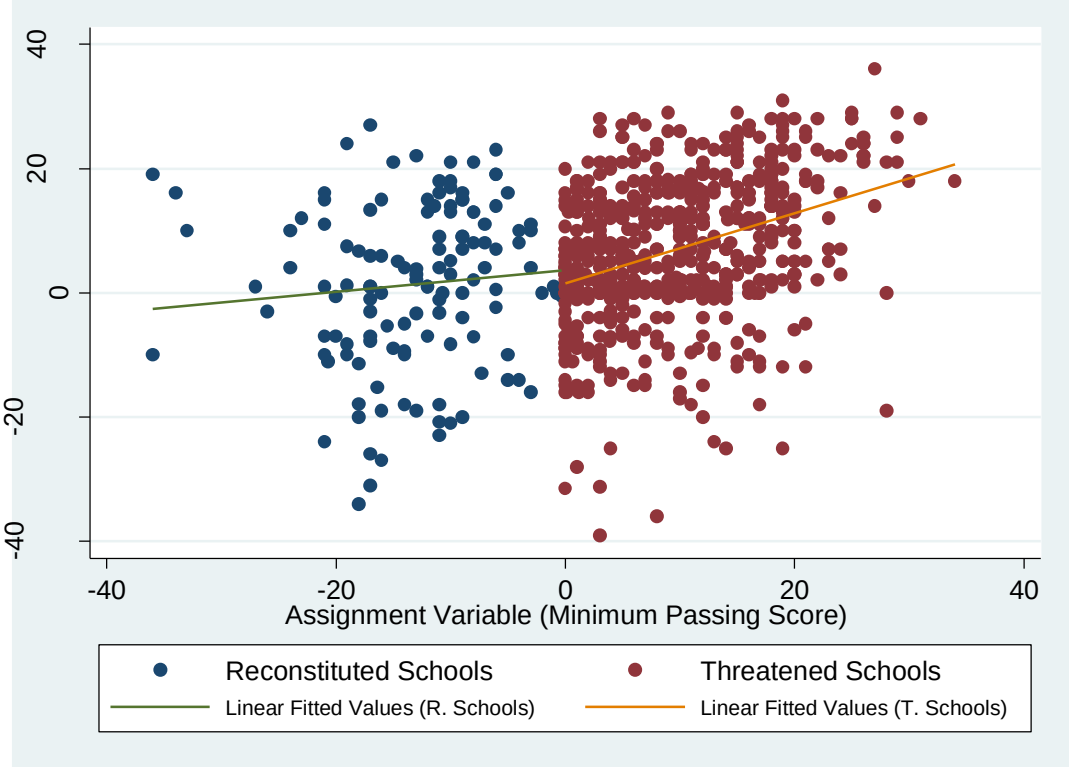

Figure 5.5 Minimum Passing Scores by Reconstitution Status

Source: AEIS Data

Average Treatment Effects

Under certain specifications of the weighting kernel (uniform) and bandwidth selection (h=40) the regression discontinuity model estimates an average treatment effect of 7.4 percentage points (Table 1), meaning that reconstitution improved school minimum passing scores.[34] Results are statistically significant for bias-corrected and robust calculations.

[34] Uniform kernel selection is intended as a benchmark estimate. More reasonable kernel selections place greater weights on units closer to the cutoff point and are presented later. All calculations in this section include 127 reconstituted schools and 675 threatened schools unless otherwise stated; the reduced sample size of reconstituted schools reflects a bandwidth selection that does not include all schools in the sample.

Method	Estimate	Std. Err.	z	P>|z|	95% Conf.	Interval
Conventional	2.0736	2.1776	0.9522	0.341	6.34152	-2.19442
Bias-corrected	7.4616	2.1776	3.4266	0.001	11.7295	3.1936
Robust	7.4616	3.2335	2.3076	0.021	13.7991	1.12405

Table 5.1 Estimates of the Effect of School Reconstitution on School Minimum Passing Rates, Standard Design[35]

Using the same model specifications as in the RD design, the RD Kink design examines whether the estimated linear functions on either side of the cutoff value experience a shift in their slope. Results shown in Table 2 provide evidence of statistically significant reduction (or flattening) in the slope of the linear function for reconstituted schools. This means, absent any RD effects, that schools with lower minimum passing scores experienced greater benefit from reconstitution than expected.

Method	Estimate	Std. Err.	z	P>|z|	95% Conf.	Interval
Conventional	0.67314	0.48212	1.3962	0.163	1.61808	-0.271807
Bias-corrected	-1.6164	0.48212	-3.3527	0.001	-0.671486	-2.56137
Robust	-1.6164	0.97009	-1.6663	0.096	0.284904	-3.51776

Table 5.2 Estimates of the Effect of School Reconstitution on School Minimum Passing Rates, Derivative (Kink) Design

[35] Results presented in Tables 5.1 and 5.2 use uniform kernel weighting and a bandwidth selection of h=40.

Sensitivity Analysis (Model Specifications)

Scholars suggest a number of methods intended to ensure results of analysis are robust to different model specifications. First, I test whether the average treatment effect relies on data farther away from the cutoff value using different 'windows' of truncated data as presented by Dong (2011). One means of accounting for potential outliers without truncating the sample involves weighting results using a kernel, often giving greater weight to data points closer to the cutoff. Average treatment effects using additional kernels are estimated and discussed. Next, the effect of bandwidth selection on estimates and their statistical significance, along with a brief discussion of parametric methods designed to optimally select the bandwidth are discussed.

Provided estimates of the effect of reconstitution on schools' minimum passing score include the entire window of potential outcome scores. Reducing the width of the window of outcomes generally reduces bias, though this might increase the variance as the sample size is reduced (Dong 2010). Table 3 provides estimates of the effect of reconstitution with various window-sized scenarios as one test of the specificity of the model with each aperture centered by the cutoff value of zero. RD estimates with the smallest window around the cutoff value do not present evidence of either a jump in reconstituted school's minimum passing score or a shift in their estimated slopes.

	RD		RD Kink		Ns
[-15,15]	-3.9253		-2.9989		78/498
	(3.6374)		(2.9913)		
[-25,25]	5.4368		-6.0971	***	121/652
	(3.5176)		(1.4807)		
[-35,35]	8.0162	**	-2.3446	**	125/675
	(3.2485)		(1.0099)		
[-45,45]	7.4616	**	-1.6164	*	127/675
	(3.2335)		(0.97009)		

Table 5.3 Average Treatment Effect of Reconstitution Using Various Apertures

Unfortunately, discontinuities rely on data removed from the cutoff value, though the relatively small sample of reconstituted schools means that smaller apertures neglect important data and reduce the sample size of both reconstituted and control schools. A non-parametric approach includes the application of statistical weighting (using a kernel) which increases the prominence of values closer to the cutoff threshold and reduces (but does not eliminate) the values farther away (Calonico 2014). Presented results so far weight school data the same regardless of their distance from the cutoff value, but two additional kernels (the triangular and the Epanechnikov) place more weight on values closer to the cutoff threshold and less weight father away (Calonico 2014). Table 5.4 demonstrates that such weighting reduces the estimated treatment effect for the RD design by roughly 2 percentage points, though effects are still statistically significant at least marginally. Kernel weighting for the RD Kink design actually increases the estimated treatment effect between the slope of the regression lines.

	RD		RD Kink	
Uniform	7.4616	**	-1.6161	*
	(3.2335)		(.97009)	
Epanechnikov	6.5519	**	-2.4025	**
	(3.0888)		(.96372)	
Triangular	5.4317	*	-2.7992	***
	(2.969)		(.93352)	

Table 5.4 Average Treatment Effect of Reconstitution, by Kernel Type

Presented statistically significant estimates rely on a relatively large bandwidth selection. Robust estimates are better able to handle relatively large bandwidth selections (Calonico 2014). Larger bandwidth selection should yield precise estimates since more data points are used, though the local linear regression may be less accurate, which might lead to bias in the estimation of the treatment effect (Jacob 2012). Figure 5.6 below calculates the share of the sample of reconstituted schools and 1-p-value (to show increasing or decreasing significance) of estimates based on using several different bandwidths. As the sample size increases beyond 85%, the p-value generally falls (shown as an increase in this graph) but only when the bandwidth selection includes more than 95% of the sample is there a statistically significant average treatment effect.

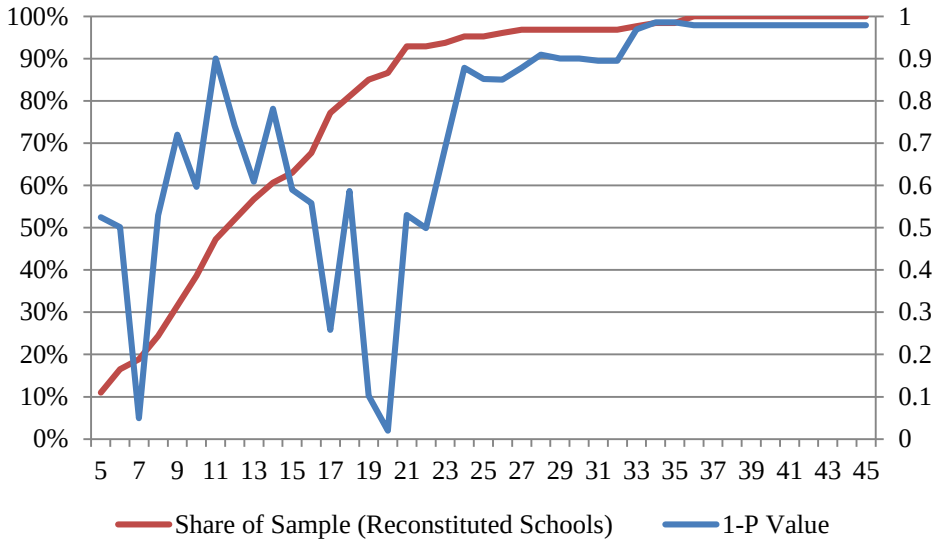

Figure 5.6 Bandwidth Selection, the Share of Reconstituted Schools Included in the RD Design, and the Statistical Significance of the Estimate of the Average Treatment Effect

Parametrically determined bandwidth selections from Calonico (2014) only include between 40% and 65% of the sample of reconstituted schools. While robust calculations applied are less sensitive to larger bandwidths than more traditional estimation methods, bandwidth sensitivity should concern any scholar (Calonico 2014). Statistically significant estimates are less reliant on bandwidth selection for the regression discontinuity Kink design; with just over 65% of the sample of reconstituted schools the RD Kink results are statistically significant.

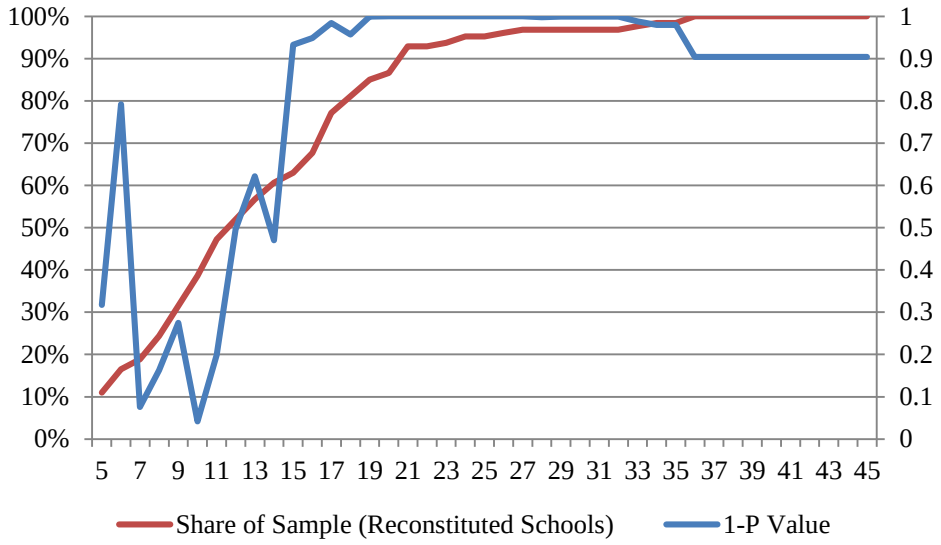

Figure 5.7 Bandwidth Selection, the Share of Reconstituted Schools Included in the RD Kink Design, and the Statistical Significance of the Estimate of the Average Treatment Effect

Sensitivities to model specifications demonstrate that average treatment effects for the RD design should be treated with some caution, though the relatively small sample size of schools hints at the need to use as large a bandwidth as possible in an effort to include all schools in the sample. However, average treatment effects for the RD Kink design are less susceptible to changes due to model specifications involving bandwidth and kernel selection.

Sensitivity Analysis (Sample Selection)

Presented results include all reconstituted schools regardless of whether they reconstituted once or more than once. However, schools required to reconstitute more than once could rightly be considered outliers. Table 5 presents average treatment effects if all reconstituted schools are included in the sample or if only schools reconstituted the first time were included in the analysis sample.

	RD		RD Kink		Ns
Including Multiple Reconstitutions	7.4616	**	-1.6161	*	127/675
	(3.2335)		(.97009)		
First Time Only	7.6041	**	-1.7055	*	124/675
	(3.2395)		(.96773)		

Table 5.5 Average Treatment Effects, With Alternate Sample Specification

Note that schools reconstituted only once appear to have higher estimated treatment effects. A relatively wide range of model and sample selections provide evidence that reconstitution improved school performance.

FIXED EFFECTS

Fixed Effects models measure the effect of long-term achievement in math and reading on students from attending the school the year it reconstituted. Estimates represent the change in the standard deviation the student test scores received in relation to students in the same grade using the same test in the same year over the course of their post-reconstitution academic career. Thus, negative numbers represent a reduction in student performance after reconstitution relative to all students in Texas while positive numbers represent an improvement in student performance after reconstitution relative to all students in Texas.

Results are broadly organized based on the sample of students included in the analysis. Since students attending reconstituted schools are always included in these specifications, the primary changes in the sample used in analysis reflect those not attending a reconstituted school. Fixed effects estimates reflect the change in student achievement relative to the comparison group. By progressively limiting the sample,

comparison group members more closely approximate the treatment group and, eventually, allow for causal interpretation of the estimates.

The first set of results use all student records in Texas, so estimates reflect changes in student achievement relative to all other students across the state. Considering this includes students at schools that never failed, positive estimates would be unexpected. The next set of results include only students who ever attended a failing or reconstituted school in Texas; since this sample is similar to that used for the regression discontinuity design above, results are expected to be positive. Finally, the sample includes students who attended a reconstituted school for the first time during its planning year (who experience reconstitution) and students who arrived at the reconstituted school in the fall after the summer it was reconstituted.

Samples of results include estimates using several different models, including OLS and various fixed effects specifications. To determine differential effects of reconstitution on student achievement for specific populations only students with these characteristics were included in the analysis; these moderating results are then presented. Next, a set of mediating variables which the literature indicates might explain or attenuate the effect of school reconstitution are progressively added. A brief summary of results are included at the end of this chapter.

All Students In Texas

Prior to examining any estimates of the effect of reconstitution in comparison to any group of students, there should be sufficient context for these results. Similar models have yet to be applied to examine school reconstitution specifically, so estimates of any effects, while interpretable, are without context. In other words, estimates of the effect of students attending reconstituted schools naturally bring up the question of the general

effect of students attending a school on their long-term student achievement in Texas. To provide this context, estimates were calculated for students attending a random selection of schools chosen over several years, as discussed in the methods chapter. The distribution of these results for Math and Reading tests are presented in Figures 5.8 and 5.9. Note that for both of these graphics, estimates include both statistically significant and non-significant results; non-significant results hover close to zero.

The range of the estimated effect of attending a school in a single year, based on assigning sets of random schools to false treatment, is from roughly -0.06 SD to + 0.05SD. The distribution of these results hint that estimates less than -0.04SD and estimates higher than 0.04SD are unlikely.

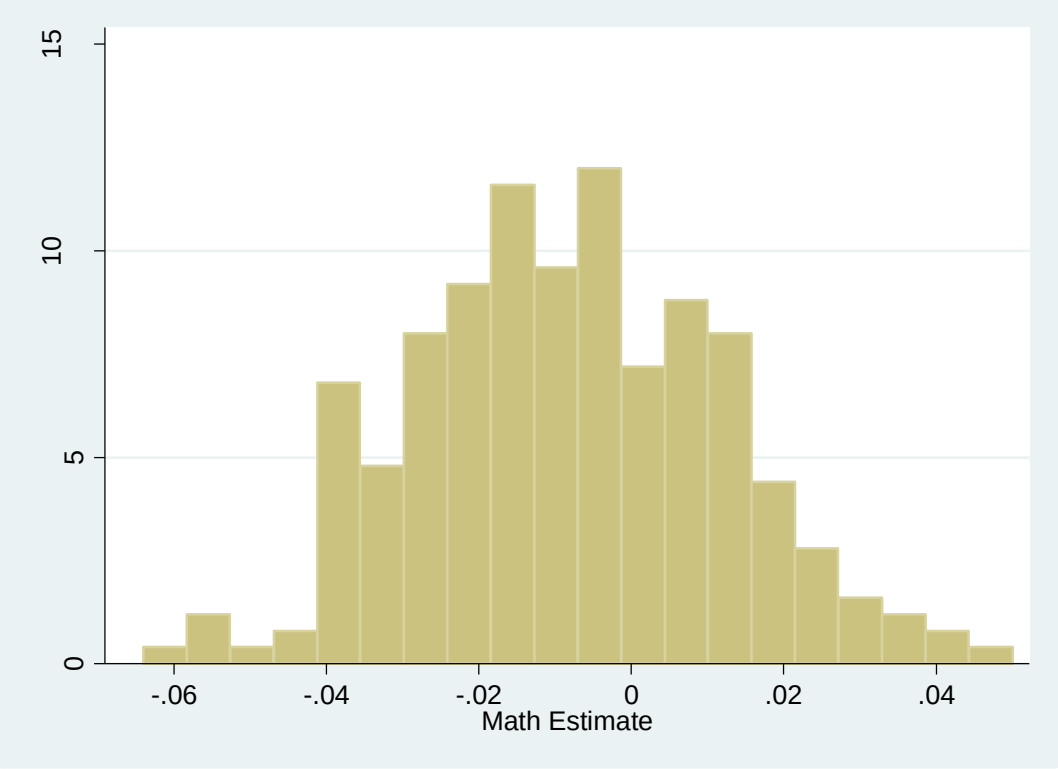

Figure 5.8 Estimates of the Change in Math Standardized Test Score Based on Attending a Randomly Selected School

Estimates of the effect of attending a school on student future reading achievement scores have a similar range, between -0.06SD though less than 0.04SD (Figure 5.9). Estimates are clustered around zero as in math estimates. For both sets of distributions, results close to zero lack statistical significance.

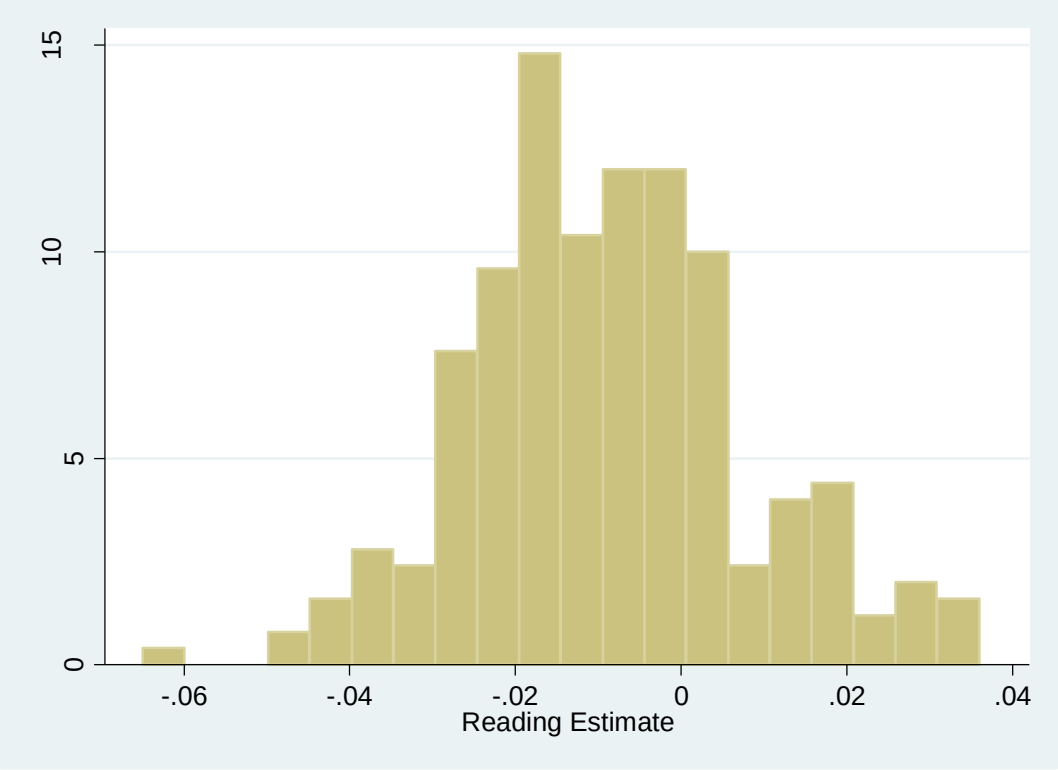

Figure 5.9 Estimates of the Change in Reading Standardized Test Score Based on Attending a Randomly Selected School

Prior to examining the effect of reconstitution on student achievement, a reasonable question is what was the effect of attending schools that were eventually reconstituted prior to reconstitution? In an effort to obtain a sense of scale regarding the state of reconstituted schools prior to their reconstitution, estimates of the effect of attending a reconstituted school the year it was classified as failing for the second consecutive time (Table 5.6). These estimates show that attending a failing school

reduced student academic achievement in math, though not in reading, relative to their peers.

	OLS		FE		FE		FE	
Math	-0.465	***	-0.030	***	-0.029	***	-0.038	***
	(0.003)		(0.003)		(0.003)		(0.003)	
Read	-0.370	***	0.005		0.005		0.005	
	(0.004)		(0.003)		(0.003)		(0.003)	
Student and Year FE			X		X		X	
Student Characteristics					X		X	
Student Years on Campus							X	

Table 5.6 Estimates of the Effect of Attending a Twice Failing School

Two years later, after reconstitution, students who attended reconstituted schools experienced a different long-term result on their student achievement (Table 5.7). Compared to their peers, students experienced a decline of a little bit more than 0.01SD in math and an increase of roughly 0.01SD in reading. While statistically significant, these results are not substantive alterations in student achievement, though student achievement in Math is not half as bad as two years ago.

	OLS		FE		FE		FE	
Math	-0.430	***	-0.012	***	-0.011	***	-0.014	**
	(0.003)		(0.003)		(0.003)		(0.003)	
Read	-0.393	***	.0076	**	.0105	**	0.012	***
	(0.004)		(0.003)		(0.003)		(0.003)	
Student and Time Fixed Effects			X		X		X	
Student Characteristics					X		X	
Student Years on Campus							X	

Table 5.7 Estimates of the Effect of Attending a Reconstituted School

One way to test the sensitivity of these results involves re-considering who received reconstitution 'treatment'. If we consider only students who 'experienced' reconstitution as those who attended reconstituted schools prior to their reconstitution, we get altered results (Table 5.8). Here, student and year fixed effects show improvements in both math and reading achievement when compared to their non-reconstituted fellow students. This hints that students experiencing reconstitution might receive greater benefit than students who attend a school after it is reconstituted. However, some of the effect of reconstitution is mediated when including the number of years a student was on campus prior to its reconstitution.

	OLS		FE		FE		FE		FE	
Math	-0.344	***	0.030	***	0.032	***	0.009	*	0.016	***
	(0.005)		(0.004)		(0.004)		(0.004)		(0.005)	
Read	-0.298	***	0.031	***	0.033	***	0.027	***	0.032	***
	(0.005)		(0.005)		(0.005)		(0.005)		(0.005)	
Student and Year Fixed Effects			X		X		X		X	
Student Characteristics					X		X		X	
Student Years on Campus							X		X	
Campus Characteristics									X	

Table 5.8 Estimates of the Effect of Reconstitution by Those Who Experienced Reconstitution

Students who experienced a school prior to and immediately after it reconstituted experienced gains in their academic standardized test scores, though for math these gains are partially attenuated by the years the student spent on the campus. When comparing results to all students in Texas for students who attended a school prior to its reconstitution, Hispanic students experience the most benefit (Table 5.9) while African American students experienced reductions in reading performance relative to other

students. Student on Free or Reduced Lunch also experienced benefits in math performance.

	White		Hispanic		Black		Asian	FRL	
Math	-0.013		0.023	***	0.002		0.082	0.020	***
	(0.014)		(0.06)		(0.008)		(0.053)	(0.005)	
Read	0.011		0.045	***	-0.030	**	0.107	0.009	
	(0.017)		(0.007)		(0.009)		(0.063)	(0.006)	

Table 5.9 Moderating Effects of Race and Economic Status for Students who Experience Reconstitution

As overall results are only marginally significant in math performance, several mediating variables appear to remove the effect of reconstitution on future student performance, including teacher and student racial and ethnic similarity, the mean number of years staff worked with the current principal, whether the school offers advanced or innovative coursework, and the share of staff new to the campus (Tables 5.10 and 5.11). Traditional measure of teacher and principal human capital, including master's degrees and tenure, strengthens the effect of reconstitution on student math performance. The estimated effect of reconstitution on reading remains generally consistent regardless of the model specification and the inclusion of mediating variables.

	No Mediating Variables		Failing Group/Failing Test		Vector Distance Between Teachers and Students		Mean Years of Staff Working with Principal at School	
Math	0.009	*	0.009	*	0.005		0.007	
	(0.004)		(0.004)		(0.004)		(0.004)	
Read	0.027	***	0.026	***	0.024	***	0.025	***
	(0.005)		(0.005)		(0.005)		(0.005)	

Table 5.10 Mediating Effects for Students who Experience Reconstitution (Part I)

	No Mediating Variables		Advanced or Innovative Coursework Offered		Traditional Measures of Teacher Human Capital		Shares of New Staff		All Mediating Variables	
Math	0.009	*	0.002		0.017	***	-0.008		-0.011	*
	(0.004)		(0.004)		(0.005)		(0.004)		(0.005)	
Read	0.027	***	0.020	***	0.032	***	0.025	***	0.027	***
	(0.005)		(0.005)		(0.005)		(0.005)		(0.005)	

Table 5.11 Mediating Effects for Students who Experience Reconstitution (Part II)

Students who Attended a Failing School

As mentioned earlier, results from above include all students on campuses, so estimates of the effect of reconstitution are compared to all students in Texas. The regression discontinuity section examines the effect of reconstitution in relation to schools that had previously failed. Since this intervention is specifically intended to improve failing schools, an additional analysis that limited the sample to include only students who attended a threatened or reconstituted school was performed; this is similar to the conditions set up for the regression discontinuity design. Students attending reconstituted schools experienced significant improvements in their standardized math and reading performance in comparison to students who attended a failing school but did not attend a reconstituted school (Table 5.12).

	OLS		FE		FE		FE	
Math	-0.168	***	0.027	***	0.029	***	0.026	***
	(0.004)		(0.003)		(0.003)		(0.003)	
Read	-0.204	***	0.021	***	0.023	***	0.023	***
	(0.004)		(0.004)		(0.004)		(0.004)	
Student and Time Fixed Effects			X		X		X	
Student Characteristics					X		X	
Student Years on Campus							X	

Table 5.12 Estimates of the Effect of Reconstitution for Students Attending a Failing School

These effects are moderated by student racial/ethnic characteristics and free and reduced lunch status (Table 5.13). White students receive no benefit or harm from attending a reconstituted school. Hispanic, Asian, and FRL students all see statistically significant improvements in their math and reading performance if they attend a reconstituted school in comparison to attending a non-reconstituted but failing school. African American students experienced real, but small benefit in their math performance, and a small negative effect on their reading performance.

	White	Hispanic		Black		Asian		FRL	
Math	0.012	0.056	***	0.017	***	0.133	***	0.043	***
	(0.012)	(0.005)		(0.007)		(0.037)		(0.004)	
Read	0.022	0.056	***	-0.018	*	0.130	**	0.026	***
	(0.013)	(0.005)		(0.007)		(0.047)		(0.004)	

Table 5.13 Moderating Effects of Race and Economic Status for Students who Ever Attended a Threatened School

Theoretically derived mediating variables fail to explain the effect of reconstitution on either math or reading performance (Tables 5.14 and 5.15). While traditional measures of principal and teacher human capital appear to strengthen the

effect of reconstitution. Estimates consistently return to around 0.026SD for math and 0.023SD for reading.

	No Mediating Variables		Failing Group/Failing Test		Vector Distance Between Teachers and Students		Mean Years of Staff Working with Principal at School		Advanced or Innovative Coursework Offered	
Math	0.026	***	0.026	***	0.023	***	0.026	***	0.019	***
	(0.003)		(0.003)		(0.003)		(0.003)		(0.003)	
Read	0.023	***	0.023	***	0.021	***	0.022	***	0.017	***
	(0.004)		(0.004)		(0.004)		(0.004)		(0.004)	

Table 5.14 Mediating Effects for Students who Ever Attended a Threatened School (Part I)

	No Mediating Variables		Traditional Measures of Teacher Human Capital		Shares of New Staff		All Mediating Variables	
Math	0.026	***	0.042	***	0.026	***	0.035	***
	(0.003)		(0.004)		(0.003)		(0.004)	
Read	0.023	***	0.035	***	0.022	***	0.0028	***
	(0.004)		(0.004)		(0.004)		(0.004)	

Table 5.15 Mediating Effects for Students who Ever Attended a Threatened School (Part II)

Special Analysis

The next analysis takes advantage of how schools reconstitute in Texas, with a planning year following failure for two consecutive years and then reconstitution in the summer after the planning year. The sample is limited to include only students who arrived on campus during the planning year and then experience reconstitution and students who arrive on campus the fall after the summer the campus reconstituted, and so

did not experience reconstitution. The smaller sample size used for this analysis allows for the inclusion of district fixed effects which serve here as a proxy for neighborhood fixed effects.[36] Regardless of the model specification below, students experiencing reconstitution experience improved performance on math and reading tests (by 0.025SD) in relation to students who attended a recently reconstituted school.

	OLS		FE		FE		FE		FE	
Math	0.121	***	0.035	***	0.033	***	0.026	***	0.025	***
	(0.008)		(0.006)		(0.006)		(0.007)		(0.007)	
Read	0.121	***	0.035	***	0.032	***	0.026	***	0.025	***
	(0.009)		(0.006)		(0.007)		(0.007)		(0.007)	
Student and Time Fixed Effects			X		X		X		X	
Student Characteristics					X		X		X	
Campus Characteristics							X		X	
District Fixed Effects									X	

Table 5.16 Estimates of the Effect of Reconstitution

Mediating variables incorporated into the last fixed effects model in Table 16 (Table 5.17 and 5.18), generally fail to attenuate the effect of reconstitution. However, some of the effect of reconstitution on reading is attenuated by adding traditional measures of human capital (teacher and principal tenure and education) to the model. Even the inclusion of all potential mediating variables only reduces the estimate of the effect of reconstitution on test performance a little.

[36] A number of additional efforts at adding various fixed effects were attempted, but computational limitations of Stata at the ERC meant these could not be calculated. These efforts include: campus fixed effects and principal fixed effects.

	No Mediating Variables		Failing Group/Failing Test		Vector Distance Between Teachers and Students		Mean Years of Staff Working with Principal at School	
Math	0.025	***	0.027	***	0.028	***	0.030	***
	(0.007)		(0.006)		(0.006)		(0.006)	
Read	0.025	***	0.025	***	0.026	***	0.026	***
	(0.007)		(0.007)		(0.007)		(0.007)	

Table 5.17 Mediating Effects of School Reconstitution (Part I)

	No Mediating Variables		Advanced or Innovative Coursework Offered		Traditional Measures of Teacher Human Capital		Shares of New Staff		All Mediating Variables	
Math	0.025	***	0.029	***	0.023	***	0.028	***	0.021	***
	(0.007)		(0.006)		(0.007)		(0.006)		(0.07)	
Read	0.025	***	0.024	***	0.023	**	0.025	***	0.022	**
	(0.007)		(0.007)		(0.007)		(0.007)		(0.005)	

Table 5.18 Mediating Effects of School Reconstitution (Part II)

CONCLUSION

Principals at reconstituted schools in Texas did not replace their teachers at rates much higher than during non-reconstitution years, though more than 40% of principals were new during the planning year and roughly 50% assistant principals were new after reconstitution. Shares of other types of staff remained relatively stable throughout the reconstitution process.

Both regression discontinuity and fixed effects designs point to statistically significant positive effects on student learning as measured by standardized test performance. Fixed Effects estimates of the effect of school reconstitution on student achievement are broadly consistent, with statistically significant effects in math (of

around 0.026SD) and in reading (around 0.023SD) for several different samples of students (those who ever attended a threatened school and those who entered a reconstituted school immediate prior to and after it reconstituted) model specifications (including control and mediating variables) and including student, year, and district fixed effects.

Chapter 6: Discussion

This dissertation aims to determine how reconstitution came to be applied as a mandatory statute in Texas in the early 2000s, what its application meant, and whether it led to improvements in school and student academic achievement, particularly for failing schools. While the number of failing schools nationwide continues to rise and reconstitution (or reconstitution-like reforms) continue to be applied, the literature on reconstitution lacks clear estimates of its effects, either positive or negative. After summarizing the results of the analysis, the rest of this chapter discusses the implications for these findings, the limitations of this work, provides suggestions for future research, and generates policy recommendations for the Texas legislature and school districts.

Reconstitution came into its own in the mid-1990s just as the accountability movement took off. It spread faster than scholars could reasonably examine it. Except for a few locations such as Chicago, schools generally reconstituted in isolation. This isolation meant that, while scholars could reasonably develop theories about how reconstitution might impact student academic performance, conclusions about student performance relied on school-level aggregated performance measures before and after reconstitution.

Reconstitution arrived in Texas statute in 2003 linked to the idea that radical reforms might bring about significant improvements, particularly for failing schools. The notion that reconstitution might improve student performance came from a nearly decade-long history of its application in Texas. However, these early experiences with reconstitution and the decision to rely on reconstitution to improve student performance at failing schools did not (and likely could not considering the small number of cases) come from a rigorous quantitative evaluation of its effects.

The original proposed language intending to place reconstitution into state statute neglected to use the term 'reconstitution,' outlining instead the specific shares of required replacement of staff. However, over several legislative sessions, the final statute excised specific staff replacement language in favor of the word reconstitution. This and the general concern in Texas of state involvement in local school district decisions led to a key difference in the application of reconstitution in Texas during the period examined in comparison to its application in other locations and times. Other locations automatically required certain shares of staff be replaced, but Texas placed the authority to keep or removed staff into the hand of principals. The evidence clearly shows that principals at these schools judiciously applied the option of removing teachers, with average rates of new teachers being similar to non-reconstitution years for the same schools, though there is some hint that teachers either chose to leave or were removed from their positons after the school failed two years in a row and prior to reconstitution.

Scholars are generally suspicious of the effects of reconstitution on student achievement. Much of the literature on reconstitution focuses on its disruptive nature. While not always stated directly by scholars, these disruptions placed the idea that reconstitution improves student achievement in doubt. A significant portion of the literature of reconstitution discusses concerns with the policy, or cautions that applying reconstitution might pose practical and legal difficulties for districts and their schools (Malen, Croninger et al. 2002, Rice and Malen 2003, Rice and Croninger 2005, Spitser 2006). This qualitative evidence hinted that any future quantitative results would demonstrate poorer outcomes for students made to suffer through the complex and confusing process of reconstitution. Additionally, locations where reconstitution ostensibly improved student achievement (such as San Francisco) received large additional funds from non-district sources to support their work. A review of the literature reasonably

led to the conclusion that situations requiring automatic reconstitution after failure without additional funding (as in Texas) would likely not improve, and could potentially harm, student achievement.

Contrary to the qualitative assessments of scholars in other locations, reconstitution in Texas appears to improve school and student academic achievement. Reconstitution improved the minimum passing score performance of failing schools in two ways. It improved the average minimum passing score by roughly 7 percentage points. It also improved the minimum passing score more for schools with the worst performance than would be expected without reconstitution. While these results are marginally subject to specifications of the regression discontinuity model, they are generally consistent.

Reconstitution improved the academic performance of students attending failing schools the years they reconstituted, though reconstitution did not improve academic performance compared to all other schools in Texas. Generally, students attending a reconstituted school improved their math performance by roughly 0.026SD and their reading performance by roughly 0.023SD. This is true for multiple samples of analysis (students attending a failing school, students who attended a reconstituted school in its planning year and the next) and regardless of model specifications (including additional control variables and, where possible, district fixed effects).

Broadly speaking, white students attending a reconstituted school received no statistically significant benefit from reconstitution. Hispanic students experienced improvements in their math (0.056SD) and reading (0.056SD) performance compared to Hispanic students attending a failing school not required to reconstitute. African American students experienced positive effects from reconstitution in math, but negative effects in reading. Asian students, whose are not singled out in the state accountability system, experienced the largest gains in performance (roughly 0.13SD for math and reading) after

attending a reconstituted school compared to their peers attending a failing school. Economically disadvantaged students also received benefits from attending a reconstituted school with improvements in reading performance (0.026SD), but particularly in math (0.043SD).

Most surprising, and perhaps frustrating, are that potential mediating variables failed to consistently attenuate the effect of reconstitution on student achievement. Campus-level teacher racial/ethnic similarity to their students (vector distance) failed to account for any improvements when comparing students attending reconstituted to failing schools. However, this similarity did meditate the effect of reconstitution in relation to all students in Texas when reconstitution is defined only as those students attending a school before and immediately after a school is reconstituted, but only for math. Additionally, the mean number of years the staff at a school worked with their principal failed to account for any reconstitution improvements in student performance. Traditional measures of human capital at schools, such as teacher and principal experience and education, tended to either increase the estimate of the effect of reconstitution (for students ever attending a threatened school) or have no effect (for students attending reconstituted schools just before and after they reconstitute). Most notably, the shares of staff new to a campus, one of the purported drivers behind the effect of school reconstitution, does not attenuate its effects on student test performance.

Results are broadly consistent across different units of analysis (school and student) and different specifications within those analyses. School-level regression discontinuity designs indicate reconstitution improves school minimum passing scores. Student fixed effects models provide evidence under multiple specifications that reconstituting schools improves student performance on math and reading tests. While results including mediating variables leave an open question about the mechanisms which might be behind

these impacts, this dissertation demonstrates that reconstitution improves student performance on standardized tests.

That leaves open several relevant questions. If established theories relating to why reconstitution might improve student achievement are true, then perhaps the measures utilized here for teacher and student racial balance or human capital insufficiently describe these characteristics. Since available data lack any link between teachers and their students save for their location, these variables may not accurately measure the interactions between student and teacher racial characteristics or teacher human capital. In this dissertation traditional measures of human capital at a school such as tenure and education suffer because they are aggregated at the campus rather than student level, hiding any classroom to classroom variation in how teachers are distributed among students. The traditional human capital measures used in this work do not include the set of vast scholarly work related to teacher 'effectiveness' related to improved student learning; were teacher and student links available in the accessed data for all the years examined such alternate measures might have attenuated the effects of reconstitution.

One implication of this dissertation is that providing principals the ability at failing schools to replace staff without requiring a certain percentage of teachers be removed produced positive impacts at Texas schools. Though given this broad authority, principals judiciously applied it. Some federal and other state efforts now ask for large shares of teachers to be removed from their campuses. While these efforts are not considered here, this dissertation demonstrates that improvements in school and student performance clearly do not require unusually large shares of teachers be removed from the school, even though such actions are currently in vogue.

Recommendations for Further Research

Several avenues of research present themselves. First, examine reconstitution on the other side of 2011, when other performance measures were in place in Texas. While outside the original scope of this dissertation, these years include some state-level linking between teachers and their students, which could allow the inclusion of alternative measures of human capital such as teacher effectiveness. This was not possible using the data described here, but could be done for a couple of years in Texas.

New state requirements passed in the most recent legislative session require campuses and their districts to submit a 'turnaround' plan to the Texas Education Agency after failing for two years in a row (HB 1842). This new law may provide additional state resources to schools (according to its fiscal note), but also allow researchers a glimpse into the planning and (ostensibly) the implementation of reconstitution-like measures.

Policy Recommendations

In an era where large administrative datasets provide the opportunity to effectively evaluate state mandated programs and policies, that such efforts rely solely on the interests (and dogged determination) of scholars means that effective policies go unnoticed and ineffective ones continue on. Reconstitution in its current form existed from 2003 to 2015 with little effort to determine whether it performed as claimed. This lack of independent scholarly oversight reduces the perception of the effectiveness of government and likely costs tax payers money. After the turn of the last century, progressives spent considerable time developing methods and means to increase the efficiency and efficacy of administration. Our current era offers a wealth of data allowing for truly accurate evaluations of programs and policies and yet too often legislative priorities and decisions neglect to ask the critical questions: "Is it possible to evaluate whether this program is working as it was intended? Is it possible to use the evaluation to improve the program's performance?" Surely Texas would benefit from inclusion in statute of a requirement to

evaluate (or at least consider the feasibility of evaluating) old and new programs and policies.

Finally, the vast wealth of administrative datasets (which allow for excellent and terrific work) would benefit from the collection of state-wide qualitative data collections. Much of the most important decisions regarding how schools and districts operate, such as the decision to use a new curriculum or the development of new tools to help counselors provide services to their students, are absent from datasets maintained in Texas. Hence, much of the variation of district choices goes unseen and unexamined. Administrative datasets are the first step, but qualitative evidence regarding what transpires on campuses should also be readily collected and maintained.

References

Academic Excellence Indicator System (AEIS), Texas Education Agency.

Gallup's Most Important Problem. G. Poll, Policy Agendas Project, Department of Government, The University of Texas.

(1954). Brown v. Board of Education. **347 U.S. 483**.

(1974). Milliken v. Bradley, US Supreme Court.

(1983). San Francisco NAACP v. San Francisco Unified School District, United States Ditrict Court, N.D. California.

(1997). Vallas urges Reconstitution of Seven High Schools. Chicago Independent Bulletin. Chicago, Ill.**:** 3.

Anagnostopoulos, D. (2003). "The New Accountability, Student Failure, and Teachers' Work in Urban High Schools." Educational Policy **17**(3): 291-316.

Anagnostopoulos, D. and S. A. Rutledge (2007). "Making Sense of School Sanctioning Policies in Urban High Schools: Charting the Depth and Drift of School and Classroom Change." Teachers College Record **109**(5).

Baum, S., J. Ma and K. Payea (2013). Education Pays 2013: The Benefits of Higher Education for Individuals and Society, CollegeBoard**:** 47.

Biegel, S. (1988). "Reassessing the Applicability of Fundamental Rights Analysis: The Fourteenths Amendment and the Shaping of Educational Policy After Kadrmas v. Dickinson Public Schools." Cornell Law Review **74**(6): 1078-1117.

BLS. (2015). "Charting the Labor Market: Data from the Current Population Survey (CPS)." Retrieved August 7, 2015, from http://www.bls.gov/web/empsit/cps_charts.pdf.

BLS. (2015). "Earnings and Unemployment Rates by Educational Attainment." Retrieved June 13, 2015, from http://www.bls.gov/emp/ep_chart_001.htm.

Calonico, S., M. Catteneo and R. Titiunik (2014). "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." Econometrica **82**: 2295-2326.

CDC (2015). Current Cigarette Smoking Among Adults in the United States, Centers for Disease Control and Prevention.

Clearinghouse, C. R. L. (2015). "Case Profile: San Francisco NAACP v. San Francisco Unified School District." Retrieved October 15, 2011, from San Francisco NAACP v. San Francisco Unified School District.

Daugherity, B. J. and C. C. Bolton (2008). With All Deliberate Speed: Implementing Brown v. Board of Education, University of Arkansas.

Dillon, S. (2009). Study Finds High Rate of Imprisonment Among Dropouts. New York Times. New York, NY.

DoE, U. (2011). Duncan Says 82 Percent of America's Schools Could "Fail" Under NCLB This Year, U.S. Department of Education.

Doherty, K. and S. Abernathy (1998). Turning Around Low-Performing Schools: A Guide for State and Local Leaders. D. o. E. P. a. E. Service, USDoE**:** 78.

Eisenhardt, K. M. (1989). "Agency Theory: An Assessment and Review." The Academy of Management Review **14**(1): 57-74.

Emery, K. (2002). "A Brief History of the San Francisco Unified School District and the Consent Decree." Retrieved June 5, 2012, from http://www.educationanddemocracy.org/SF/Brief_history_SF.htm.

Emery, K. (2002). The Business Roundtable and Systemic Reform: How Corporate-Engineered High-Stakes Testing has Eliminated Community Participation in Developing Educational Goals and Policies, University of California Davis.

Gardner, D. P. (1997). "A Nation At Risk." Congressional Digest **76**(11): 258.

Goldstein, J., M. Kelemen and W. S. Koski (1998). Reconstitution in Theory and Practice: The Experience of San Francisco. Annual Meeting of the American Educational Research Association. San Diego**:** 47.

Hamilton, M. P., J. V. Heilig and B. L. Pazey (2014). "A Nostrum of School Reform? Turning Around Reconstituted Urban Texas High Schools." Urban Education **49**(2): 182-215.

Hendrie, C. (1997). "Reconstitution gaining new momentum." Education Week **17**(4): 1.

Hess, G. A. (2003). "Reconstitution—Three Years Later: Monitoring the Effect of Sanctions on Chicago High Schools." Education and Urban Society **35**(3): 300-327.

Hoffmann, J. P., L. D. Erickson and K. R. Spence (2013). "MODELING THE ASSOCIATION BETWEEN ACADEMIC ACHIEVEMENT AND DELINQUENCY: AN APPLICATION OF INTERACTIONAL THEORY." Criminology **51**(3): 629-660.

Imbens, G. W. and T. Lemieux (2008). "Regression discontinuity designs: A guide to practice." Journal of Econometrics **142**(2): 615-635.

Kingdon, J. W. (1984). Agendas, Alternatives, and Public Policies. New York, NY, Harper Collins.

Kluger, R. (2004). Simple Justice: The History of Brown v. Board of Education and Black America's Struggles for Equality. New Your, New York, Knopf.

Koury, E. T. (2000). Reconstitution: The principal's view. 9968309 Ed.D., Harvard University.

Lee, D. S. and T. Lemieux (2010). "Regression Discontinuity Designs in Economics." Journal of Economic Literature **48**(2): 281-355.

Lewis, W. D. and T. V. Young (2013). "The Politics of Accountability: Teacher Education Policy." Educational Policy **27**(2): 190-216.

Malen, B., R. Croninger, D. Muncey and D. Redmond-Jones (2002). "Reconstituting Schools: "Testing" the "Theory of Action"." Educational Evaluation and Policy Analysis **24**(2): 113-132.

McCrary, J. (2008). "Manipulation of the running variable in the regression discontinuity design: A density test." Journal of Econometrics **142**(2): 698-714.

Miller, M. and B. Montgomery (1976). Teacher Integration. Los Angeles Times (1923-Current File). Los Angeles, Calif.**:** 1-c4.

Orfield, G. (1999). "Policy and equity: A third of a century of educational reforms in the United States." Prospects **29**(4): 579-594.

Parker, W. (2008). "Desegregating Teachers." Washington University Law Review **86**(1): 53.

Phillips, D. C. (2014). Behaviorism. Encyclopedia of Educational Theory and Philosophy. D. C. Phillips. Los Angeles, SAGE Reference. **1:** 79-83.

Retana, N. M. (1997). Radical school restructuring: A case study of principal leadership in reconstituted urban high schools. 9825062 Ph.D., The University of Texas at Austin.

Rice, J. K. and R. G. Croninger (2005). "Resource Generation, Reallocation, or Depletion: An Analysis of the Impact of Reconstitution on School Capacity." Leadership and Policy in Schools **4**(2): 73-103.

Rice, J. K. and B. Malen (2003). "The Human Costs of Education Reform: The Case of School Reconstitution." Educational Administration Quarterly **39**(5): 635-666.

Richardson, V. (1997). Constuctivist Teacher Education: Building a World of New Understandings. Brisstol, PA, The Palmer Press.

Rojas, W. (1996). Reconstitution, Reculturing, and Reform: Adding Options for Urban Education. PhD Dissertation, Columbia University.

Ronfeldt, M., S. Loeb and J. Wyckoff (2013). "How Teacher Turnover Harms Student Achievement." American Educational Research Journal **50**(1): 4-36.

Rozmus, K. (1998). "Education Reform and Education Quality: Is Reconstitution the Answer?" BYU Education & Law Journal **1998**: 103-145.

Ruiz-de-Velasco, J. L. (1999). The politics of education in court-ordered school districts: School reform in San Francisco. 9924493 Ph.D., Stanford University.

Rylander, C. (1999). Challenging the Status Quo. T. Comptroller. Austin, TX.

Soledad, A. M. (2006). Critical factors related to successful school reconstitution in Texas. 3214007 Ed.D., The University of Texas at El Paso.

Spitser, A. (2006). "School Reconstitution under No Child Left behind: Why School Officials Should Think Twice Comment." UCLA L. Rev. **54**: 1339-1384.

SPLC. (2004). "Brown v. Board: Timeline of School Integration in the U.S." Teaching Tolerance Retrieved May 15, 2015, from http://www.tolerance.org/magazine/number-25-spring-2004/feature/brown-v-board-timeline-school-integration-us.

Usher (2011). AYP Results for 2011-2012. Washington, D.C.

Usher (2011). How Many Schools Have Not Made Adequate Yearly Progress?, Center on Education Policy**:** 11.

Wong, K. K., D. Anagnostopoulos, S. Rutledge, L. Lynn and R. Dreeben (1999). Implementation of an Educational Accountability Agenda: Integrated Governance in the Chicago Public Schools Enters Its Fourth Year**:** 49.

Zavadsky, H. (2012). School Turnarounds: The Essential Role of Districts. Cambridge, Mass., Harvard Education Press.